ROYAL AIRCRAFT ESTABLISHMENT
MAIN LIBRARY
ALDERSHOT 24461 EXT. 954
AUTHOR
N.A.T.O. - A.G.A.R.D.
UDC 621.396.969.1
COPY A
DATE 14.10.66.

AGARDograph 100

RADAR TECHNIQUES FOR DETECTION TRACKING AND NAVIGATION

RADAR TECHNIQUES FOR DETECTION TRACKING AND NAVIGATION

Proceedings of the Eighth Symposium of the
AGARD Avionics Panel
London, 21–25 September 1964

Edited by

W. T. BLACKBAND

Royal Aircraft Establishment, Farnborough, England

Published for and on behalf of

ADVISORY GROUP FOR
AEROSPACE RESEARCH AND DEVELOPMENT
NORTH ATLANTIC TREATY ORGANISATION

by

GORDON AND BREACH SCIENCE PUBLISHERS

NEW YORK · LONDON · PARIS

Published by

GORDON AND BREACH SCIENCE PUBLISHERS INC.

150 Fifth Avenue, New York, N.Y. 10011

Distributed in the United Kingdom by:

BLACKIE AND SON LTD.

5 Fitzhardinge Street
London W.1., England

Distributed in France by:

DUNOD EDITEUR

92 rue Bonaparte
Paris 6, France

Distributed in Canada by:

THE RYERSON PRESS

299 Queen Street West
Toronto 2B, Ontario, Canada

Copyright © 1966

ADVISORY GROUP FOR
AEROSPACE RESEARCH AND DEVELOPMENT
NORTH ATLANTIC TREATY ORGANISATION

Library of Congress Catalog Card Number 65-23648

Printed in Northern Ireland at

THE UNIVERSITIES PRESS, BELFAST

FOREWORD

The Avionics Panel of AGARD held its eighth annual Symposium in London from 21st to 24th September, 1964. This meeting was devoted to a study and review of the current radar techniques for the detection and tracking of targets, and their application to navigation. The history of radar is that of the modern world. In its various stages it has been limited by the technicological limits of the day, reacted upon these limits so as to stimulate new branches of technology, and then bounded forward to further sophistication. The growth of radar has not only been limited by technology, political restraints have been equally effective, indeed it is so difficult to trace the early development because in the various countries in the early 1930's radar either languished through government indifference, or was hidden by a cloak of secrecy. However, the political forces arising from military needs having their effect, each of the major powers had its indigenous radar by the time of World War II.

If radar made its most sensational growth for military purposes its roots were in the pacific study of the ionosphere. It is interesting to recall that although the basic pulse method of ionosphenic sounding was first published by Breit and Tuve in 1926, the radio measurements of ionosphenic height previously published by Appleton and Barnett had used an FM transmission. The pulse method displaced the FM method largely because of the relative ease of processing the measured data. In fact one of the chief factors governing the development of modern radar has been the problem of processing and handling the incoming information.

The development of the transistor has transformed electrical methods of computing, and consequently whereas in 1939 conversion from polar to Cartesian co-ordinates and other data processing was done by equipment built around Post Office uni-selectors and relays, today there seem to be few limits to the processing complexity which we can hope to employ. Much of the first 19 chapters of this book, in which the design of radar systems is discussed, reflects the effect of the expanding computing capabilities upon the planning of the present generation of radars and in particular upon the choice of their modulation systems. In parallel with this are the chapters 28–32 which discuss primarily the newer methods of handling data.

The development of radar has ever stimulated the associated branches of engineering and the procedure devised for the design of the close tolerance parabolic reflector for the "Haystack" installation described in Chapter 20

is likely to have a wide application to other structural engineering projects—perhaps as unexpected as any of the by products of our study of radar.

It is strange that this new branch of engineering which first flowered in war should now benefit so many civil projects. We now expect aircraft and ships to have radar guidance and we have radar to enforce speed limits on our roads, it has become part of our everyday life. But not content with this, radar now stretches out into space, and dares to wrestle with the weather a truly Protean adversary of everchanging form, but one capable of revealing truth as the reward of persistence and determination.

W. T. BLACKBAND

CONTENTS

	Foreword	v
Chapter 1:	A New Philosophy of Radar—*M. H. Carpentier*	1
Chapter 2:	Parametric Analysis of Spaceborne Radar Applications—*R. J. Orford*	9
Chapter 3:	Analysis of Unco-operative Radar Targets—*R. M. Goldstein*	21
Chapter 4:	Pulse Compression Research in the United Kingdom—*E. H. Boyenval*	35
Chapter 5:	Méthodes Statistiques en Radiodétection—*L. Reboul*	47
Chapter 6:	Target Resolution: Capabilities of Modern Radar and Fundamental Limits—*A. Rihaczek*	75
Chapter 7:	Measurements of the Reflection from the Ground of S-band Signals—*A. R. Domville*	97
Chapter 8:	Some Information on Nature, Size-and Velocity Distribution of Random Scatterers gathered by a Pulse Radar on two Wavelengths using Vertical Fixed Beams—*W. Fogy*	111
Chapter 9:	Moving Target Indication: A Survey of Developments since 1948—*P. Bradsell*	141
Chapter 10:	Le Radar SDS, Radar de Surveillance Destiné à Détecter les Objectifs Mobiles au Sol:extension à la Détection des Avions Volant Bas—*G. van den Broek*	173
Chapter 11:	Sideways Looking Airborne Radar—*J. Highcock*	185
Chapter 12:	Airborne Early Warning Radar: Some Basic Considerations—*S. Matt*	203
Chapter 13:	Airborne Doppler Navigator Techniques—*T. Gray*	217
Chapter 14:	Scanning Beam Guidance for Approach and Landing—*J. E. Woodward*	235
Chapter 15:	An Automatic F.M.C.W. Surveillance Radar—*C. S. E. Phillips*	253

Chapter 16: Doppler System for Measuring the Flight Performance of Small Rockets and Light Ammunition—*C. O. Lund* 267

Chapter 17: Integrated Trajectory System—*G. F. Bigelow* 281

Chapter 18: Methods for obtaining Velocity and Range Information from CW Radars—*M. Easterling* 307

Chapter 19: A Description of a High-performance Microwave Experimental Facility—*H. G. Weiss* 327

Chapter 20: The Design of a very High Power, very Low Noise Cassegrain Feed System for a Planetary Radar—*P. D. Potter* 373

Chapter 21: Electronically Scanned Antenna Systems, I—*M. A. Diab and T. Maggio* 397

Chapter 22: Electronically Scanned Antenna Systems, II—*E. A. Killick and D. E. N. Davies* 417

Chapter 23: High Power CW Radar Transmitter—*W. S. Baumgartner* 433

Chapter 24: Transmitter Research in the U.K.—*N. S. Nicholls* 467

Chapter 25: Phase Shifters for Phased Array Antennas—*H. A. Hair* 473

Chapter 26: Dispersors for Pulse Compression Systems—*P. S. Brandon, O. E. Keall and W. S. Mortley* 485

Chapter 27: Radar Data Handling and Display Systems for use with Pulse Radars—*A. P. Young* 515

Chapter 28: Digitalisation of Radar Signals and Their Evaluation by a Computer for Automatic Tracking of Targets—*K. v. Schlachta, H. Springer, W. Storz and W. D. Wirth* 525

Chapter 29: Evaluation of Track-While-Scan Computer Logics—*L. Leth-Espensen* 541

Chapter 30: Tracking in an Air Traffic Control Environment—*B. Oakley* 565

Chapter 31: Provision and use of Alphanumeric Display of Primary and Secondary Radar Tracking Data in Expanding Air Traffic Control Systems Capacity—*L. E. Shoemaker* 585

Index of Names 601

Index of Places and Institutions 605

Index of Subjects 607

CHAPTER 1

A NEW PHILOSOPHY OF RADAR

MICHEL H. CARPENTIER
C. F. T. H, Bagneax (Seine), France

Conventional radars are based upon pulse train modulation, and they measure range in terms of pulse delay. The measurement of this may be made by the use of either correlators or matched filters. The choice of transmitter modulation is discussed. This choice should take into account the modulation effects of a scanning antenna. The receiver should be matched to the signal reflected from the target, which for large moving targets may differ appreciably from the transmitted signal.

1. CONVENTIONAL RADAR—INHERENT LIMITATIONS

For many years almost every radar has been designed in accordance with the same concepts, these are:

— the transmission of a rectangular pulse (in fact, a train of pulses) during which the microwave frequency and phase remain as nearly constant as possible;
— the measurement of the delay between the start of the transmitted signal and the start of the received signal returned from a target.

This received signal is, on the one hand, of a very low level, and on the other hand, accompanied by a noise emanating from either the radar receiver itself or from enemy jamming where military equipment is concerned (or even from jamming caused by an awkward friendly source).

It is, therefore, a signal identical to the transmitted signal but reduced in amplitude and shifted in time, to which a noise is added. The following two remarks should be made straightaway:

Firstly, the signal which is received is not in fact identical to the transmitted signal: this could only be so if the target were small compared with the wavelength, or if the target were a metal sphere. The fact that the target has non-negligible dimensions gives rise to a received signal the amplitude and phase of which are not constant during the duration of the signal, the form of the signal received varying, for instance due to the relative movement of the target in relation to the radar.

Secondly, the manufacturer of radar systems must ensure that the spectrum density of the noise be reduced as much as possible by using low noise receivers and antennae of low noise temperature.

In the past with radars making the approximation that the received signal from a useful target was not modulated in phase and/or in amplitude within the pulse the reception was ensured by filtering through a narrow band filter having a bandwidth equal to the reverse factor of the transmitted pulse duration. (This kind of radar, was not the only one used since certain altimeters measured aircraft height by the transmission of a frequency-modulated signal, but such equipments were not considered really as radars.)

When it is necessary to increase the range of a conventional pulse radar and that the utmost has been done as regards the antenna and the reduction of the interference noise spectrum density, there are only two solutions:

(a) increase the peak power, that is to say, the pulse power; the limits possible or reasonable at the present time have been reached;
(b) reduce the noise power by reducing the receiver bandwidth, which necessitates the increase of the transmitted pulse duration and deteriorates the quality of the range measurement, because in a conventional radar the range measurement becomes better as the transmitted pulse becomes shorter.

At the present time the requirements for radar fall into three categories:

(i) radars with very extensive range (some thousands of kilometers) having good range measurement qualities for very small targets;
(ii) medium-range radars having a very good range measurement quality when in the presence of considerable noise due to jamming;
(iii) short-range radars having extremely accurate characteristics.

These requirements could only be met with conventional radars by the transmission of very high powers.

II. THE IDEAL RECEIVER

Having reached this point, it is necessary to revert to the past. What has been the main reason for the existence of conventional radars? It seems that if the signal transmitted by the radar is a signal modulated in amplitude only, it is because the state of the art made available to technicians magnetrons capable of transmitting considerable power during very short periods but incapable of frequency modulation according to a defined law. But now amplifier tubes capable of transmitting any type of signal are available and there is no longer any reason to be restricted to amplitude modulation.

Consequently, if we can choose freely, the character of the transmitted signal $S(t)$, and if we suppose that the signal received from a target is identical to the transmitted signal apart from an attenuation and a delay, it can then be shown that the best chance of detecting a target is obtained by making the correlation of the received signal $Y(t)$ with a signal identical

to the transmitted signal delayed through the range of all possible delays, that is to say, the calculation of:

$$C(t_0) = \int_T Y(t) \cdot S(t - t_0)\, dt,$$

T being the useful duration of the transmitted signal.

But it is also shown that exactly the same result is obtained by having the received signal pass through a filter "matched" to the transmitted signal. This is demonstrated thanks to a theorem expounded by mathematicians which is referred to as "the convolution theorem."

III. SELECTION OF THE TYPE OF RECEIVER

Thus, at least as long as another analogous theorem has not been found (and it does not seem that such be possible) there are two bases for the design of a radar receiver, i.e.:

— the use of correlators, and
— the use of matched filters.

When one uses correlators one is dealing with correlation radars, of which the "Pulse Doppler" constitutes the oldest type. When a matched filter is employed, one is dealing with matched filter radars, of which the conventional-type radar constitutes the oldest, and the "CHIRP" pulse compression radars one of the most recent forms.

The choice between these two types of radar depends on technological considerations:

If the transmitted signal is composed of a random signal modulated in phase, it is generally impossible to manufacture a matched filter and so correlators are used.

If the transmitted signal is composed of pulses during which the frequency varies in a linear manner with time, matched filters are more economical than correlators are.

If the transmitted signal is composed of a train of pulses not modulated in phase, both types of receivers are, *a priori*, possible and the choice between them depends on the precise value of the transmission parameters as well as practical considerations.

IV. CHOICE OF THE SIGNAL TO BE TRANSMITTED

But if the choice of the receiver depends on considerations relevant to the status of the art and on the nature of the transmitted signal, how can the modulation of the transmitted signal be chosen?

The choice depends on three considerations, i.e.

(1) The necessary range, which does not determine the power but the energy to be transmitted during the measurement duration, that is, energy which is often more easy to obtain with a large measurement duration and a low power than with a tremendous power during a very short time.

However the measurement duration is limited to a maximum which depends either on the maximum time which can be waited for to obtain the measurement result or the time during which the parameters to be measured are sufficiently constant taking into account the required accuracy for the measurement, or the time during which it can be considered that the received signal does not fluctuate too much, taking into account the behaviour of the target and any movement of the radar antenna.

(2) The required accuracy for radial range measurements and radial velocity which precludes certain signals having too narrow a spectrum (inaccurate radial range measurement) or too short a duration (inaccurate radial velocity measurement) and certain signals giving rise to ambiguous measurements.

(3) The strength, radial range and radial velocity of targets other than the target of interest.

Let us suppose that a certain transmission signal has been selected: $S(t)$. In conjunction with this signal there is what is called an ambiguity function, i.e.

$$\mathscr{A}^2(\theta\,,F)$$

depending on the difference θ between the radial range of a foreign target and that of the target concerned, and the difference F between their radial velocities.

This function represents the degree of disturbance brought about by this interfering target during the detection and localization of the target concerned.

Thus, knowing the geography of the foreign targets (i.e. clutter) represented by the following function:

$$V(\theta\,,F)$$

given the "strength" of the clutter in relation to θ and to F, the disturbance brought about by these foreign targets is represented by the integral:

$$\iint \mathscr{A}^2(\theta\,,F)\cdot V\,(\theta\,,F)\,d\theta\,dF.$$

Thus, $\mathscr{A}^2(\theta\,,F)$ must be chosen, and therefore $S(t)$, in order that this integral be as low as possible.

One, has, therefore, a criterion for the selection of the optimum modulation of the transmitted signal, after which the type of receiver to be utilized can be chosen.

Remark

This selection has not in fact always been made correctly, because it has been forgotten that, in the direction of the target, the transmitted signal is the result of the modulation by the antenna of the signal at transmitter output.

For example with a panoramic radar the antenna movement introduces an amplitude modulation of the signal transmitted by the whole of the equipment and the receiver must not be matched to the signal form at transmitter output but to the form of the signal received by the target, which is quite a different matter.

V. MATCHING THE RADAR WITH THE TARGET

In other respects, it does not suffice to select the transmitted signal $S(t)$ in order to minimize the disturbance provoked by clutter, clutter response, that is to say, integral:

$$\iint \mathscr{A}^2(\theta, F) \cdot V(\theta, F)\, d\theta\, dF$$

but $S(t)$ should also be chosen to maximize the useful signal response, that is to say, the same expression in which $V(\theta, F)$ represents the geography of the useful target.

This consideration is not important when,

(a) the dimensions of the useful target are small compared with the radar range resolution;
(b) or the Doppler spectrum of the useful target is narrow in comparison to the Doppler frequency resolution of the radar, which has often been the case up to now.

But it is becoming very important to take into account this consideration when

(c) the target is very extended in range and velocity, in cases when, for example, the useful target is the ground or the moon;
(d) or the resolving power in range of the radar becomes very good.

This concept leads to a further evolution in radar. Past radars have been based upon the assumption that the signal reflected by the target has the same form as the transmitted signal, which is not true.

In fact, the signal emanating from the target is the transmitted signal modulated by the target and what is necessary to bring about is to employ a receiver matched not to the transmitted signal but to the signal received when $S(t)$ is transmitted, and when one is dealing with a specified target.

The next evolution in radar will probably be the utilization of a receiver matched to such a target when it is illuminated by a given transmitted signal by means of a specified antenna, thereby obtaining the possibility of:

(i) obtaining a very much improved response,
(ii) obtaining a distinct response only from a specific form of target. Radars equipped with several outputs each corresponding to a given type of target can thereby be conceived.

CONCLUSION

It can be seen at this stage that one is far removed from the empiricism of conventional pulse radars and that present-day materializations are still far from having exhausted discussions on the "radar subject" in which technicians will still have plenty of time for the constructive examination of the matter.

DISCUSSION

R. W. WILMER: You mentioned a practical experiment in which you achieved a 20 db gain by matching your received signal to the target characteristics rather than to the transmitted signal. Can you please tell us more about this?

M. H. CARPENTIER: We were able to have a large gain by matching the radar receiver not to the transmitted signal but to the received signal, that is to say the transmitted signal transformed by the target. This experiment was made on the Moon and in actual fact in matching the characteristics of the receiver to the very special form of the signal returned by the Moon we were able to obtain a gain of 20 db in the quality of the reception.

J. MURPHY: You mention that the main theme of your paper was that perhaps a radar system should be designed on the basis of the obstacle which it is to detect. In this regard, in the light of Mr. Wilmer's question I would like to know if you have any thoughts on the detection of wire-like objects by a radar system that must perform other functions such as ground mapping.

M. H. CARPENTIER: At a given moment you usually are interested in one particular target, and if you choose a transmission signal appropriate to the nature of the target and make a receiver adapted to the signal as transformed by the target you can, on the one hand, considerably increase the conformity of the detection to the signal and at the same time decrease the received signals from any targets with which you are not concerned at that moment.

E. RECHTIN: If I understand the integral expression which you have derived, you have given us a criterion for obtaining an optimum transmitted signal but we still do not have the process by which you can tell by some automatic algebra what the best modulation is. Am I correct?

M. H. CARPENTIER: In fact the only appropriate process we have as far as I know is to choose by commonsense, a number of signals which apparently are likely to give a minimum or a maximum integral and solve for these using a computer. As far as I know there is no general procedure which would make it possible by taking into account a given clutter-geography and to find automatically the best signal to transmit.

R. J. LEES: The transmitted signal must be decided by the degree of resolutions you want to get, i.e. the band width you must explore in order to get resolution in range and the phase shift you must explore to get resolution in phase . . . . Having determined that you then optimise your reception circuit for getting the information from the source. Is this not so?

M. H. CARPENTIER: Yes, there are three considerations which help in

the choice of the transmitting signal. First the duration of transmission, secondly the characteristics of the range and radial velocity measurements, and thirdly that you have to avoid ambiguous measurements that arise from clutter of foreign targets. The first two considerations eliminate a certain number of signals that don't conform, and you have to take this consideration of clutter to help you to determine your choice.

CHAPTER 2

PARAMETRIC ANALYSIS OF SPACEBORNE RADAR APPLICATIONS

R. J. ORFORD
TRW Space Technology Laboratories, Redondo Beach, California, U.S.A.

In this paper we have shown the significant relationships between radar parameters and mission parameters pertaining to search and acquisition. The most important of these is the linear variation of radar power-aperture product with detection range, closing velocity and search area, and the optimization of search frame time. The linear variation with range was unexpected and indicates a tradeoff between detection range and spacecraft propulsion requirements. Curves are presented for numerical analysis of situations with Swerling Case I type targets. Analysis of other target types showed that the increased cross-section fluctuation may demand a considerably higher signal to noise ratio based on the average cross-section. On the other hand, due to the fluctuations, the average cross-section was larger than expected so that the overall radar power-aperture product for detection might not be appreciably affected.

1. INTRODUCTION

The problem of searching for and acquiring a satellite target from another satellite or missile is one of wide interest in both civilian and military space technology. In this paper we shall be concerned with those missions when the target is non-cooperative either because of malfunction or intent. Of particular interest is a rescue mission wherein a friendly craft has become inoperative. We shall analyse the search and acquisition problem to relate radar parameters (power, aperture, etc.) to mission parameters (detection range, closing velocity, etc.) illustrating certain optimizations that can be made.

Consider a spacecraft approaching and seeking to acquire another satellite as a target. As a result of guidance errors in our own spacecraft accrued through the launch and midcourse phases, and uncertainties in the target satellite location due to imperfection in groundbased tracking data, the precise coordinates of the target satellite relative to our own spacecraft are unknown. From a model of our own guidance system and of the groundbased tracking system, it is possible, however, to ascertain a best estimate of the target's relative position and velocity with an associated distribution of error. From the mission requirements and the positional uncertainty, a volume inside which the target lies (with appropriate probability) can be defined. Let us call this volume the target uncertainty volume.

For most cases of practical interest, in terms of the relative geometry between the two craft, we can safely ignore the orbital nature of the situation for this phase of the mission. That is, we need only consider the rectilinear relative motion of the two craft, independent of the effect of gravitational attraction of a nearby body, viz. the earth, as this enters only in differential terms. Therefore, the problem posed to our rescue acquisition system is to search for the target in a specified volume in space (relative to our craft) which is approaching at a constant velocity. This velocity (vector) is also specified along with an uncertainty from the aforementioned best estimate and error distribution. We shall assume herein that this best estimate velocity (vector) is colinear with the best estimate of target position. This is a reasonable simplifying assumption and merely specifies that the guidance in the preceding phases would cause a successful intersection of the most probable target location. Of course, offsets or misses are sometimes purposely introduced because of mission requirements, however, to the extent that these usually occur, the modification to the ensuing analysis, while tedious, has small effect.

2. SYSTEM ANALYSIS

During acquisition, the acquisition radar scans (searches) over the volume in which the target is contained, repeating the scan pattern until a detection is indicated. A number of scans is usually possible and it is therefore the cumulative probability of detection P_c that we are interested in as distinguished from the single scan detection probability or so-called blip-scan ratio, P_d. The basic problem is to determine the minimum radar transmitted power and aperture required to provide the desired cumulative probability of detection (P_c) at the specified detection range (R_d) while not exceeding a specified probability of false alarm (P_{fa}).

The detection probability will depend on the radar output signal to noise ratio. Assuming coherent integration of the signal, this signal to noise is given in terms of radar parameters by

$$(S/N) = \frac{P_t G A \sigma T_i}{(4\pi)^2 k T_e L R^4} \tag{1}$$

P_t = average transmitted power
G = transmitting antenna gain
A = receiving antenna effective area
σ = target radar cross section
T_i = integration time (or time on target)
k = Boltzmann's constant
T_e = effective noise temperature of the receiver
L = system losses
R = range to target

The assumption of coherent integration will be evaluated later.

Now consider introducing the effects of mission parameters into Eq. (1). First of all, the antenna gain is equal to

$$G = \frac{4\pi}{\omega}$$

where ω is the solid angle of the radar beam. Next, assuming the system dwells a constant time on target throughout the scan,

$$T_i = T\frac{\omega}{\Omega}$$

where

T = search scan frame time
Ω = search solid angle

Furthermore, with the target closing at velocity V, we can define

$$\Delta = VT$$

as the distance travelled by the target in a search frame time.

Substituting these expressions into Eq. (1), we obtain

$$(S/N) = \frac{P_t A\sigma\Delta}{4\pi k T_e L V \Omega R^4} \tag{2}$$

Now consider the search solid angle Ω. We shall assume that this angle is fixed so as to enclose the target uncertainty volume at the specified detection range R_d. (Actually this angle can be programmed as the target closes so as to provide improved power density at the target at long range, however, this improvement will be left for future analysis.) In terms of the cross sectional area A_s of the target uncertainty volume as viewed from our spacecraft,

$$\Omega = \frac{A_s}{R_d{}^2}$$

Substituting into Eq. (2) and rearranging we find that at R_d the required radar power aperture product is

$$P_t A = 4\pi k T_e L \frac{V A_s R_d{}^2}{\sigma\Delta}(S/N) \tag{3}$$

To this point neither (S/N) or Δ have been prescribed.

Consider the minimization of Eq. (3). The object of the radar design is to achieve a certain cumulative probability of detection P_c at the range R_d with specified target. On each scan, the radar will pass over the target and obtain a particular (single-look) probability of detection P_d. A choice exists between

(a) a slow scan dwelling a relatively long period of time on the target to achieve a high P_d per scan (by signal integration), but thereby being able to achieve only a few scans as the target closes
(b) a fast scan dwelling a relatively short period of time on target, sacrificing P_d per scan for the improvement possible through the cumulative effect of many scans.

Since P_d is a function of (S/N), then it is a function of R, viz. $P_d(R)$. The above choice therefore indicates that the cumulative probability is a function of Δ as well, viz. $P_c(R, \Delta)$. Now the cumulative probability of detecting the target when it has reached a particular range R' is

$$P_c(R', \Delta) = 1 - \prod_{l=0}^{m} [1 - P_d(R' + l\Delta)]$$

assuming scan to scan independence. Here m is set by $P_d\ (R' + m\Delta) \rightarrow 0$ or by initiation of the search phase. The final scan on all targets must take place such that $R_d \leq R' < R_d + \Delta$, so that the overall cumulative probability of detection relative to R_d is just

$$P_c(R_d, \Delta) = \frac{1}{\Delta} \int_{R_d}^{R_d+\Delta} P_c(R', \Delta)\, dR'$$

As might be expected, there is an optimum scan frame increment Δ, or rather the ratio (Δ/R_d). This optimization has been addressed by Mallett and Brennan[1], Dishington[2] and Alexander et al[3]. This optimum ratio turns out to be a function of the desired cumulative detection probability P_c. Since P_d is a function of (S/N), the signal to noise (S/N) at R_d is also prescribed in the optimization procedure, so that Eq. (3) can be written

$$P_t A = 4kT_e L \frac{VA_s R_d}{\sigma} C(P_c) \tag{4}$$

where

$$C(P_c) = \left. \frac{\frac{S}{N}}{\frac{\Delta}{R}} \right|_{R_e}$$

To determine $C(P_c)$ the blip-scan relationship $P_d(R)$ corresponding to the target under consideration must be specified. A plot of the quantity $C(P_c)$ vs. Δ/R_d is presented in Fig. 1 for a slowly fluctuating target (Swerling Case I). The minimum of various curves represents the optimum values of $C(P_c)$. We should note that these values are for coherent integration and for a particular false alarm rate. The effects of these parameters will be discussed shortly.

PARAMETRIC RESULTS

With Eq. (4) and the numerical results of Fig. 1, we may determine the requisite radar power aperture product in terms of mission parameters. From Eq. (4) we see the following important results: the radar aperture product varies

(a) linearly with the target closing velocity V

(b) linearly with the cross-sectional area A_s of the target uncertainty volume as viewed from the searching vehicle

and

(c) *linearly* with the required detection range R_d.

This linear variation of power-aperture product with range is indeed unexpected and extremely significant. Conventionally, we expect a higher order (fourth power) dependence placing a premium on detection range in the overall system. However, the derived linear dependence suggests that the overall system design take advantage of increasing detection range to reduce terminal guidance propulsion requirements and improve guidance accuracy. An optimization of the overall system design in terms of detection range, trading off radar weight with propulsion, can lead to surprisingly long ranges.

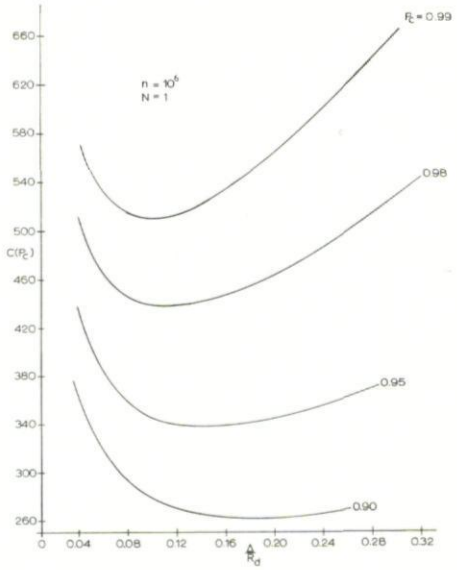


FIG. 1. *Normalised power-aperture as a function of scan interval for a slowly fading target (Swerling Case I).*

On the other hand, the effect of closing velocity is clearly detrimental, increasing both radar power-aperture, vehicle propulsion requirements and reducing the available terminal guidance time.

Similarly, the effect of the target uncertainty volume is solely detrimental. The linear dimension of this volume represents possible miss that must be accommodated by terminal phase propulsion whereas the cross section of the volume directly effects the radar. Therefore, reducing guidance and target tracking errors prior to acquisition has a significant effect on the overall system design.

Example

As an illustrative example of the optimum (minimum) radar power aperture product consider detecting with 0.99 probability at 100 km a 2 m^2 target closing at a rate of 500 m/sec. The target has been specified to lie in a sphere of 20 km radius. Let the losses be 10 db and assume a receiver

noise temperature of 1000°K. From Fig. 1, the minimum (optimum) value of $C(P_c)$ is 510. This optimum occurs at a frame ratio (Δ/R_d) corresponding to a scan frame time of 20 seconds, which seems reasonable although the specific antenna requirements should be evaluated. Using the above parameters, we obtain a power aperture product of

$$P_t A = 7 \text{ watt-m}^2$$

For example, with an antenna 0.5 m in diameter and an antenna efficiency of 0.5, the average transmitted power required is approximately 70 watts. Bear in mind that this is a minimum value and must be modified to account for non-coherent data processing and the other effects which might increase the power requirements as much as 5 or 6 db.

EFFECT OF NON-COHERENT INTEGRATION AND FALSE ALARM TIME

The previous derivation has assumed the radar output to be coherently processed. This may not be practicable in a particular case for a number of reasons and the implication of non-coherent processing must be considered further.

The important ingredients to this aspect of the problem are the duration of the target return (T_i) and the bandwidth (B) of the system immediately preceding a (the) detector. It has been established[4] that the output of a detector driven by noise of bandwidth B can be adequately represented as a series of independent noise variates of duration $(1/B)$. Coherency implies that the target return consist of one noise variate or that

$$BT_i = 1$$

If this is not the case, then post-detection (or non-coherent) integration is required and the number of noise variates N will be simply

$$N = BT_i$$

It must be emphasized that the number of noise variates N (or noise pulses) does not correspond to the number of radar transmitter pulses except in special cases. This point is discussed further in Bussgang et al.[5] and Marcum[6].

The false alarm probability P_{fa} depends on the total number of noise variates that occur in the entire radar detection system in the specified false alarm time T_{fa}. If the radar has m channels then this number of (false alarm) variates is

$$n = mBT_{fa}$$

and is termed the false alarm number.

For a slowly fluctuating target Swerling[7] has shown that the probability of detection (blip-scan ratio) can be approximately represented by the following functional form

$$P_d \cong \exp\left\{-\frac{g(n, N)}{(S/N)}\right\}$$

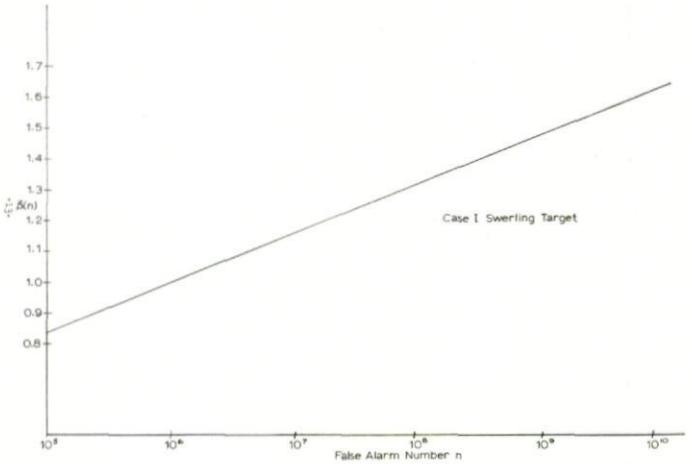


FIG. 2. *False alarm correction factor for a slowly fading target.*

Figure 1 was obtained for $(n = 10^6, N = 1)$. For other values of (n, N) we simply must alter $C(P_c)$ or

$$C(P_c) = C(P_c, n, N) = \frac{g(n, N)}{g(10^6, 1)} C(P_c, n = 10^6, N = 1)$$

For systems analysis, $\dfrac{g(n, N)}{g(10^6, 1)}$ can be adequately represented by

$$\frac{g(n, N)}{g(10^6, 1)} = \beta(n)\, \gamma(N)$$

These functions are presented in Figs. 2 and 3 as derived from the curves in Swerling[7].

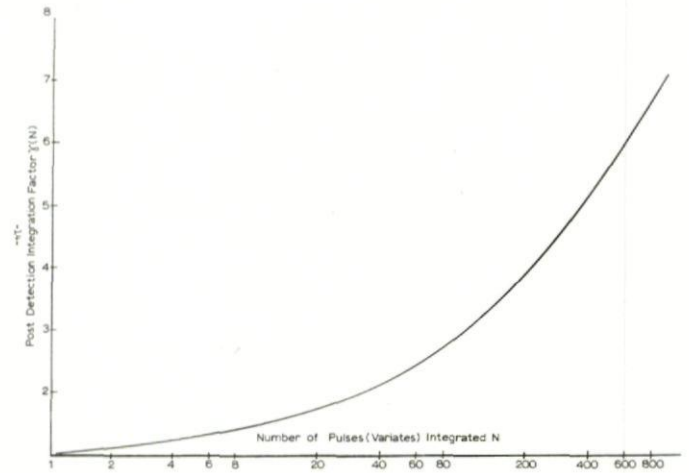


FIG. 3. *Post detection integration loss factor for a slowly fading target.*

One of the major factors contributing to the difficulty of achieving coherent detection is the rate of change of Doppler shift caused by target motion, i.e. during the time on target T_i, the frequency of the returning signal from the target will vary due to changes in the relative velocity of the two vehicles. Let θ be the angle between the line of sight (LOS) to the target and the relative velocity vector $\bar{V}$ (see Fig. 4). The Doppler shifted

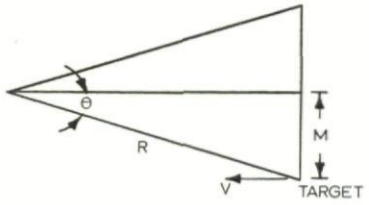


FIG. 4. *Typical (one-dimensional) geometry.*

return signal will have frequency

$$f = f_0 + \frac{2V}{\lambda} \cos \theta$$

where f_0 is the carrier frequency and λ is the nominal wavelength. The rate of change of signal frequency will be

$$\dot{f} + \frac{2}{\lambda} (\dot{V} \cos \theta - V \sin \theta \, \dot{\theta}) \tag{5}$$

However,

$$\dot{V} \cos \theta = -\ddot{R} = \text{acceleration along the LOS}$$

Also, with M as the lateral displacement of the target,

$$\sin \theta = \frac{M}{R}$$

so that,

$$\dot{\theta} = \frac{MV}{R^2}$$

Substituting, into Eq. (5) we obtain

$$\dot{f} = \frac{2}{\lambda} \left(\frac{(M^2V^2}{R^3} - \ddot{R} \right)$$

Now the bandwidth of the predetection circuitry must be wide enough to allow for this shift at the maximum (M, V) and as close as R_d, i.e.

$$B \geq \dot{f} \, T_i$$

The prediction bandwidth must also be wide enough to accommodate the spectrum of the carrier.

The time on target is determined by the dimensions of the antenna beam and the scan rate. To a reasonable approximation

$$T_i = T\frac{\omega}{\Omega}$$

where

ω = antenna beam volume

Ω = search volume

The number of noise variates N that must be integrated post detection is then

$$N = BT_i \geq \dot{f}T_i^2$$

Since N involves $\dot{f}$ and T_i, which in turn involves $(M, R, V, \ddot{R}, \lambda)$, additional effects of the mission parameters can be seen. Usually specific trends can be determined only by numerical analysis as these additional effects do not have any reasonable analytic form. However, the general effect of wavelength λ can be readily seen. Since the time on target T_i depends on the beam volume ω, then T_i is proportional to λ^2. The bandwidth B is either at some minimum value (thereby independent of λ) or equal to $\dot{f}T_i$ (thereby proportional to λ). Therefore N is proportional to λ^2 or λ^3 showing the desirability of short wavelengths ($N = 1$ is optimum).

5. OTHER TARGET MODELS

It is of interest to note that there are only three radar cross-section distributions in general use; viz, constant and Swerling Case I and Case III. While these are reasonable models, they appear to have been selected principally because of the tractability of analytically obtaining their associated detection probabilities. It seems appropriate to present and illustrate here a technique for obtaining the detection probabilities for an arbitrary distribution, particularly one obtained experimentally. The probability of detecting a signal is a function of the signal to noise at the detector output which, in turn, is a linear function of the target cross-section, i.e. from Eq. (1)

$$\left(\frac{S}{N}\right) = f(R, P_t, \ldots)\sigma$$

The appropriate measure of radar detection[7,8] is simply the expected value of the detection probability, viz, letting $\left(\frac{\bar{S}}{N}\right)$ be the signal to noise for the average cross-section $\bar{\sigma}$, then

$$P_d\left(\frac{\bar{S}}{N}\right) = \int_0^\infty P_{d_0}\left(\frac{S}{N}\right) p\left(\frac{S}{N}\middle|\frac{\bar{S}}{N}\right) d\left(\frac{S}{N}\right) \tag{6}$$

where

$$p\left(\frac{S}{N}\middle|\frac{\bar{S}}{N}\right) = \left(\frac{\bar{S}}{N}\right)^{-1} p\left(\frac{\sigma}{\bar{\sigma}}\right)$$

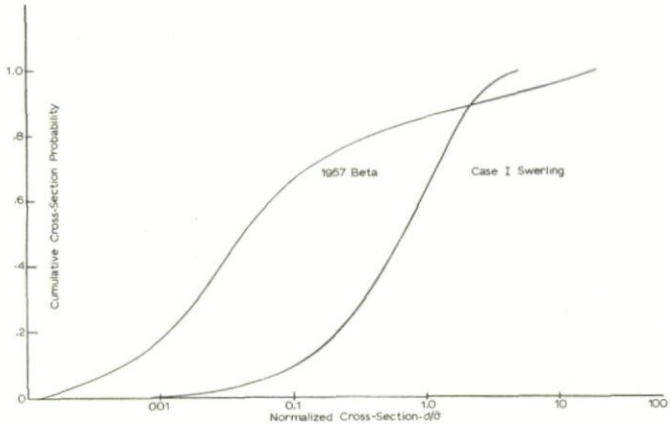


FIG. 5. *Cumulative cross-section probability for fluctuating targets.*

The function P_{d_0} is the constant signal blip-scan ratio obtained from Marcum[6]. $p(\sigma/\bar{\sigma})$ is the probability density of the cross-section normalized to the mean and must be obtained from experimental data or otherwise specified. The integration of Eq. (6) is performed numerically. As an illustration of this technique, the cross-section data for satellite 1957 Beta (from reference 9) shown in Fig. 5 was analysed, with the results shown in Fig. 6. Shown for comparison in both figures are the corresponding distribution and detection

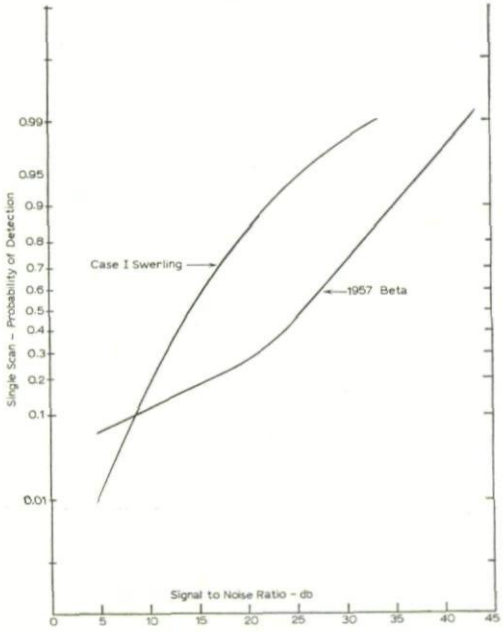


FIG. 6. *Probability of detection for fluctuating targets.*

characteristics for a Swerling Case I target of the same mean value of cross-section. The difference in detection characteristics should be of considerable interest to a system designer. The large increase in (S/N) required is associated with the high probability that $\sigma/\bar{\sigma} < 1$.

It should be noted that due to the large cross-section fluctuations, the mean cross-section $\bar{\sigma}$ for the data of Fig. 5 was nearly 40 m^2 which is quite large compared to the physical cross-section of 1957 Beta. Therefore it is possible that this target would be no more difficult, or perhaps even easier, to detect than a similarly sized target following a Swerling Case I fluctuation.

Optimization of the scan frame ratio for this target distribution model can be carried out using these results directly in the methods described previously[1,2,3].

DISCUSSION

C. BELL: I wonder if it is possible to go back to Dr. Orford's paper this morning and ask him who or what was "Swerling." In going to coffee break I found out a lot of my friends who usually know everything didn't know this so I am now emboldened to show my ignorance and ask. There have been so many long standing differences between Europe and the United States on standard targets and standard echoing areas and this appears to be a little bit out of somebody's code or series of standard targets and it is jolly important that Europe should know what these are and how they are defined.

R. J. ORFORD: Peter Swerling was a member of the Rand Corporation, and was concerned in much of the original work in the United States on detection theory. He derived two targets distribution, Swirling Case I, which is exponential and Swirling Case III which is the Rayleigh distribution.

REFERENCES

1. MALLETT, J. D. and BRENNAN, L. E., Cumulative probability of detection for targets approaching a uniformly scanning search radar, *Proc. IEEE*, **51** (April 1963).
2. DISHINGTON, R. H., Radar detection, Internal Memorandum, TRW Space Technology Laboratories, 29 Aug. 1963.
3. ALEXANDER, F., HARRIS, J and ORFORD, R., Corrections to Mallett and Brennan curves of cumulative probability of detection, Internal Memorandum 9323.1-155, TRW Space Technology Laboratories, 23 Mar. 1964.
4. MAXIMON, L. C. and RUINA, J. P., Some statistical properties of signal plus narrow band noise integrated over a finite time interval, *J. Appl. Phys.*, **27,** 1442–1448; 1956.
5. BUSSGANG, J. J., NESBEDA, P. and SAFRAN, H., A unified analysis of range performance of CW pulse, and pulse Doppler radar, *Proc. Inst. Radio Engrs., N.Y* pp. 1753–1762 (October 1949).
6. MARCUM, J. I., A statistical theory of target detection by pulsed radar, *IRE PGIT*, April 1960.
7. SWERLING, P., Probability of detection for fluctuating targets, Research Memorandum 1217, RAND Corporation, or *IRE PGIT*, April 1960.
8. SCHWARTZ, MISCHA, Effects of signal fluctuation on the detection of pulse signals in noise, *Trans IRE* **IT-** 66–71 (June 1956).
9. SKOLNIK, M., *Introduction to Radar Systems*, McGraw Hill, 1962.

CHAPTER 3

THE ANALYSIS OF UNCOOPERATIVE RADAR TARGETS

R. M. GOLDSTEIN
Jet Propulsion Laboratory, Pasadena, California, U.S.A.

By far the least cooperative class of targets that is likely to be accessible to radar is the group of our distant neighbors, the planets. This lack of cooperation stems from the all but incredible distances involved. However, when the modern high-power transmitter, large antenna, and ultrasensitive receiver are teamed together, these targets are detectable. In fact, echoes from Venus, Mars, Mercury, and Jupiter have been detected at JPL's DSIF station at Goldstone.

When the received power is large enough so that there is some margin left over from the requirements of straight detection, it becomes possible to gather much information about the target. Each point of the target has two attributes that can be measured independently at the receiver: its round-trip distance (or time delay), and its velocity (or Doppler shift). Consider a rough, rotating planet as a target. All points which have the same time of flight lie on a circle, so that dividing the received power according to time delay is equivalent to dividing the target into concentric rings, centered about the sub-Earth point. All points which have the same Doppler shift (due to rotational velocity) are found to project as straight lines parallel to the axis of rotation. Thus dividing the received power according to frequency shift is equivalent to dividing the target into parallel strips.

This paper discusses the special modulation and detection processes which enable the target to be analyzed into these two dimensions simultaneously. Such a system was used with Venus as the target during the conjunction of 1962, and a much improved version will be in operation for the conjunction of June 1964. It is expected that this system will yield very accurate measurements of the range of Venus, as well as the rotation rate, direction of the axis of rotation, and the detection of surface features on Venus.

1. INTRODUCTION

As radar targets, the Moon and the planets are most uncooperative, a fact which stems from the tremendous distances involved. First contacts with such a target were made in 1946 and 1947, when a group from the U.S. Signal Corps[1] and a group in Hungary[2] succeeded in detecting radar echoes from the Moon.

A more difficult astronomical target is Venus; even at times of closest approach, it is 10 million times less cooperative than the Moon. Our

technology required only 15 years of developing higher power transmitters, antennas of greater collecting area and focusing power, and more sensitive receivers in order to span this seven orders of magnitude. Radar contact with Venus was first made during the inferior conjunction of the spring of 1961 by four groups, working independently: JPL[3] and MIT[4] in the United States; Jodrell Bank[5] in England, and a Crimean station[6] in the USSR.

At the far point of its orbit, Venus becomes more difficult as a radar target by another factor of 1000. Our technology has now covered this factor, and echoes have also been detected from Mars, Mercury, and Jupiter.

The basic task of radar astronomy is the analysis of such targets, i.e., the utilization of radar echoes to determine, insofar as is possible, the physical characteristics of the target planet. The greatest difficulty in accomplishing this task is the almost incredibly minute amount of power contained in a planetary echo. For example, the typical power density of a Venus echo, when that planet is at the far point of its orbit, is only 3×10^{-25} watts (0.2 photons per sec) per square meter. Because of background noise, which is eternally present, signal processing is a constant struggle to conserve signal-to-noise ratio.

2. TOTAL POWER

The measurement of power is fundamental to this task. If the signal is strong enough, more can be learned about the target if the echo power is divided according to its frequency content or according to its distribution in time. However, after such partitioning, the measurement reduces to the determination of the amount of power which remains in the selected "cell."

Total power measurements give directly the radar cross section of the target planet. Surface characteristics can be related to the cross section, and target motion and rotation can be related to time variation and periodicities of the cross section.

Power is defined as the average of the square of signal, so a natural method of power measurement follows this definition. Figure 1 is a block diagram of this method. The filter limits the amount of noise which is presented to the square law device.

The output w, on the average, equals the sum of the noise power through the filter w_n and the signal power w_s

$$\bar{w} = w_s + w_n \tag{1}$$

where the bar denotes ensemble average. Because of the random nature of the signal and the noise, w varies about its average value with a standard deviation of σ_w.

$$\sigma_w^2 = \bar{w}^2 - \bar{w}^2$$

A performance criterion SNR (or postdetection signal-to-noise ratio) is

$$\text{SNR} = w_s/\sigma_w \tag{2}$$

The design problem is to make SNR as large as possible. The following equations hold:

$$u(t) = \int_{-\infty}^{\infty} h(x)\,[s(t-x) + n(t-x)]\,dx$$

$$v(t) = u^2(t)$$

$$w = \frac{1}{T}\int_0^T v(t)\,dt$$

It can be shown that, in order to maximize SNR, $H(f)$ must be chosen such that

$$|H(f)|^2 = S_s(f) \tag{3}$$

where $S_s(f)$ is the power spectral density of the signal, and the noise has been assumed white. It can also be shown that, under the assumption of white Gaussian noise, the block diagram of Fig. 1 results in an optimum

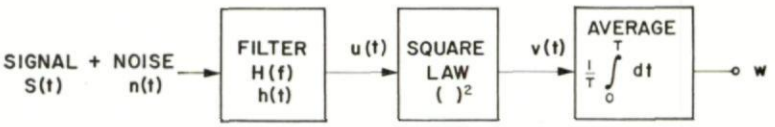


FIG. 1. *Power meter block diagram.*

power meter (radiometer). When $H(f)$ is properly chosen, the SNR that results is

$$\text{SNR} = \frac{1}{N}\sqrt{T\int_0^{\infty} S_s^2\,(f)\,df} \tag{4}$$

where T is the integration (observation) time and N is the noise spectral density (one-sided).

Equation (4) contains two inexorable laws of radar: for a given signal power, those signals which have larger $\int S_s^2(f)df$ (i.e., narrow-band signals) are more easily detected, and a factor of k loss in signal power can be compensated for by a factor of k^2 increase in observation time.

There are two difficulties in applying Eqs. (3) and (4) to a practical situation. First is the fact that $S_s(f)$ is usually unknown. Indeed, it may well be the quantity we set out to measure. Therefore, it is often assumed that $S_s(f)$ is a rectangular function of some bandwidth B. Of course, a search problem exists in the case of the weaker signals. They may not be detectable at all unless the proper bandwidth is utilized.

When such an $S_s(f)$ is substituted into Eq. (4), it becomes

$$\text{SNR} = \frac{w_s}{N}\sqrt{\frac{T}{B}} \tag{6}$$

which is the well-known radiometer formula.

The second difficulty is the problem of resolving w into its two parts, w_s and w_n. There are always troublesome gain variations in a practical

system which make this separation difficult. One common method of solving this problem is to key the signal on and off (easily done at the transmitter) and identify w_s, with the *change* in w. Thus the measurement is independent (to first order) of varying system gain. Of course, this technique reduces SNR somewhat,

$$\text{SNR} = \frac{w_s}{2N}\sqrt{\frac{T}{B}} \tag{7}$$

Equation (7) is known as the switched radiometer formula, and the factor of $\frac{1}{2}$ is the cost of that practical consideration.

3. POWER DISTRIBUTION IN FREQUENCY

When the received signal contains more power than the requirements of pure detection, it becomes possible to learn much more about the character-

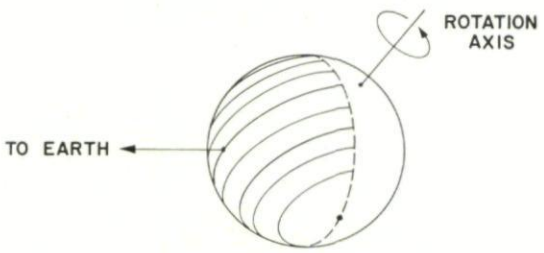


FIG. 2. *Contours of constant frequency shift.*

istics of the target planet by measuring the power spectrum of the echo. For this measurement a spectrally pure sine wave is transmitted. The echo is both shifted and broadened in frequency by the Doppler effect. The shift is caused by the relative orbital velocity between the planet and the radar station. Measurements of this shift have been used to compute the astronomical unit[7] and will be used to improve the Venus and Earth ephemerides.

The Doppler broadening is caused by any apparent rotation the planet may have, since different parts of the surface will have different line-of-sight velocities. Rotation has two components: one is the normal spin about an axis, and the other is the relative orbital motion, which also produces Doppler broadening. The situation is illustrated in Fig. 2, where the contours of constant line-of-sight velocity (and hence frequency shift) have been drawn. These contours are circles, such that they appear to the radar station as equispaced lines, all parallel to the projected axis of rotation.

Thus a spectrogram of the echo power is equivalent to scanning the disk with a narrow slit parallel to the contours of Fig. 2. The amount of frequency

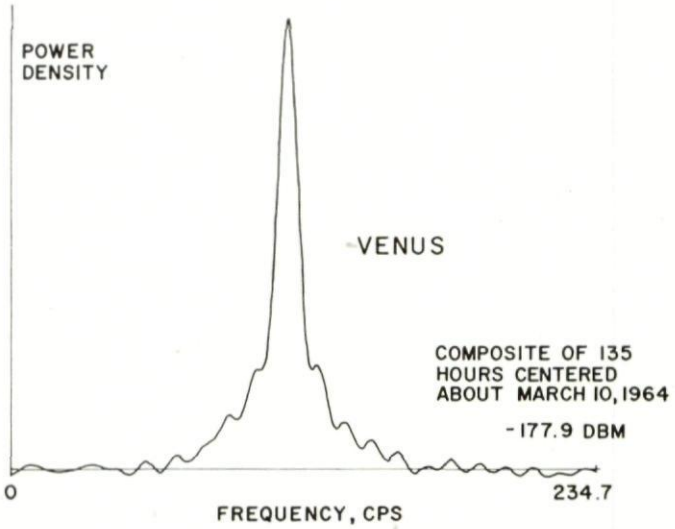


FIG. 3. *Venus spectrogram.*

broadening B, corresponding to reflections from the extreme contours, is

$$B = \frac{4\Omega r}{\lambda} \tag{8}$$

where Ω is the projected rotation rate, r is the radius, and λ is the wavelength. Observations of Venus during the conjunction of the fall of 1962[8] produced the surprising results showing that Venus rotates *retrograde*, with a period of about 250 days.

Inferences about the scattering properties of the surface can be made, based upon the shape of the spectrum. For example, if the surface is smooth, most of the power will be returned from the sub-Earth area, and the spectrum will be highly peaked. Figure 3 is a sample spectrogram taken of Venus at a distance of 88 million miles from Earth. To be quantitative, we define the backscatter function $F(\theta)$ to be the radar cross section of a unit area of surface, taken as a function of the angle of incidence. Figure 4 illustrates this point.

Given $F(\theta)$ and the assumptions of a homogeneous surface (all elements of the sphere have the same $F(\theta)$), and a surface such that the power from

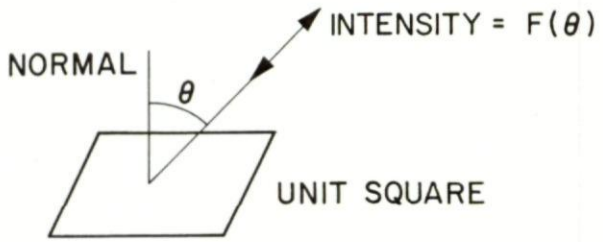


FIG. 4. *Backscatter function.*

individual elements adds, then the signal spectrum $S_s(f)$ can be computed as

$$S_s(f) = \int_{\sin^{-1} f/a}^{\pi/2} \frac{F(\theta) \sin \theta \, d\theta}{(a^2 \sin^2 \theta - f^2)^{1/2}} \tag{9}$$

where a is a rotation constant.

Equation (9) shows the dependence of $S_s(f)$ upon $F(\theta)$. This equation has been inverted to yield

$$F(\theta) = -\frac{2a^2 \cos \theta}{\pi} \int_{a \sin \theta}^{a} \frac{S'(f) \, df}{(f^2 - a^2 \sin^2 \theta)^{1/2}} \tag{10}$$

Equations (9) and (10) form a transform pair. Thus measurement of the spectrum of the echo reveals the back-scattering characteristic of the surface.

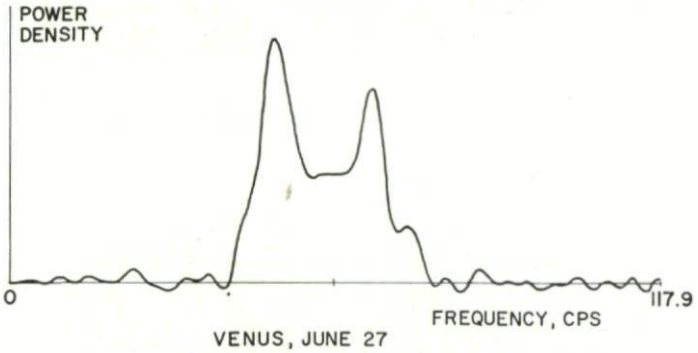


FIG. 5. *Cross-polarized Venus spectrogram, June 27, 1964*

In radar astronomy, transmission is normally circularly polarized. A single reflection reverses the sense of circular polarization, so the receiver is usually set to receive the opposite sense. However, if the surface is rough compared to the size of the wavelength, double reflections will occur. If the transmitter and receiver are both circularly polarized in the same sense, only double reflections (or any *even* number) will be received. Figure 5 is a sample spectrogram taken of Venus in the light of double reflections only.

The total power in this mode is less by a factor of about 18 (for Venus), but interesting features are revealed. The very strong central spike is missing and certain rough areas of Venus are clearly seen. As Venus rotates, these areas move slowly across the spectrograms, from the high-frequency side to the low.

The spectrograms shown in the figures were made by the autocorrelation approach[9]. Figure 6 is a block diagram of the system. The maser, programmable local oscillator, and several stages of IF amplification and conversion, etc., are all labelled BANDPASS FILTER. Samples x_i of the signal are taken at the Nyquist rate. They are quantized to only two levels by the limiter. The correlator forms the following sums of the quantized signal y_i:

$$R_y(k) = \frac{1}{I} \sum_{i=1}^{I} y_i y_{i+k} \qquad k = 1, 2, \cdots, K \tag{11}$$

where $R_y(k)$ is the estimate of the autocorrelation function of the signal at y. Two-level quantization is done to make the sums of Eq. (11) easy to mechanize. The y_i always equal ± 1, and multiplication and addition are greatly simplified.

One might expect the action of the limiter to alter the spectrum beyond recovery; but fortunately this is not so, and a simple formula relates the autocorrelation function at y to that at x.

$$R_x(k) = R_x(0) \sin\left[\frac{\pi}{2} R_y(k)\right] \tag{12}$$

This correction is made by the computer, which then calculates the spectrum by:

$$S(f) = 2 \sum_{k=1}^{K} R_x(k) \cos(\pi k f / W) + 1 \tag{13}$$

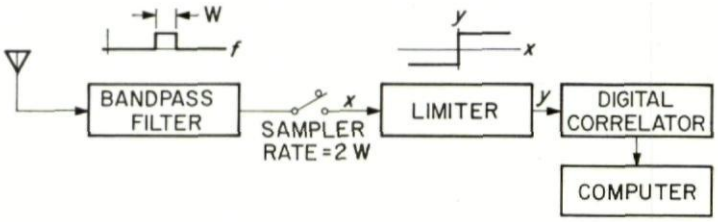


FIG. 6. *Spectrometer block diagram.*

It can be shown that this method is equivalent in performance to a set of stagger-tuned, optimum radiometers, except for a factor of $2/\pi$. That is the cost in performance of the simplification allowed by use of only two-level quantization.

In practice, a measured spectrum is always the sum of the signal spectrum plus the noise spectrum. It is desirable not only to subtract off the noise spectrum but to estimate the total power as well in order to provide a measurement of the radar cross section. The raw material for this is two spectra: $P(f)$, the sum of signal plus noise, and $Q(f)$, the noise only. Typical spectra such as these are illustrated in Fig. 7. They have been normalized to unit power to remove the problem of system gain variations.

$$P(f) = \frac{S_s(f) + S_n(f)}{w_s + w_n}$$

$$Q(f) = \frac{S_n(f)}{w_n} \tag{14}$$

A good estimate $W(f)$ of the signal spectrum is

$$W(f) = P(f) - kQ(f) \tag{15}$$

where k is a constant to be determined. If k is properly chosen, the average of $W(f)$ over the frequency interval C (where there the signal component is known to be zero) would be zero. This condition defines k:

$$k = \frac{\int_C P(f)\, df}{\int_C Q(f)\, df} \tag{16}$$

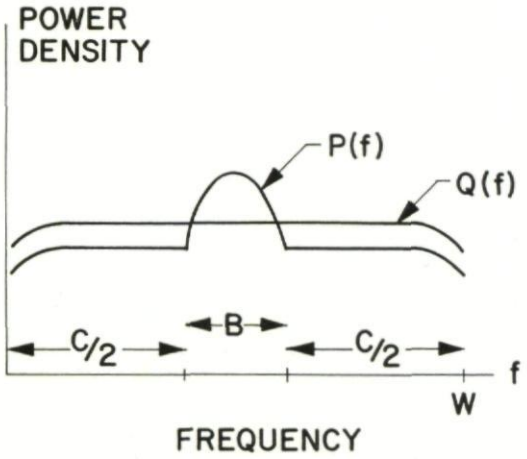


FIG. 7. *Signal-plus-noise and noise-only spectra.*

It can be seen, from Eqs. (14) and (16) that

$$w_s = w_n \frac{1 - k}{k} \tag{17}$$

The performance criterion SNR of this method of determining the total power is

$$\text{SNR} = \frac{w_s}{N}\sqrt{\frac{CT}{2(W - C)\, W}} \tag{18}$$

where W is the total spectrometer bandwidth. Equation (18), for $C = W/2 = B$,the signal bandwidth, is exactly the switched radiometer formula (Eq. 7).

In the limit as $C = W - B$ and $W \to \infty$ (of course, the spectrometer complexity increases), Eq. (18) becomes

$$\text{SNR} = \frac{w_s}{N}\sqrt{\frac{T}{2B}} \tag{19}$$

It must be noted that measuring $P(f)$ and $Q(f)$ requires twice as much time as measuring $P(f)$ only: but, in the monostatic radar case, this time is available while the transmitter is on.

4. POWER DISTRIBUTION IN RANGE

Range measurement is essentially a problem of estimation of time delay. A modulated signal is transmitted, and the time required for the signal to return is a direct measure of the distance.

Detection theory tells us that (under some simplifying assumption) the best way to estimate time delay is to compare, by correlation, the signal with the possible expected signals, of various time delay. The closest comparison corresponds to the best estimate of time delay. Even where the simplifying assumptions are not strictly met, it is always good strategy to compare the actual signal with the possible expected ones.

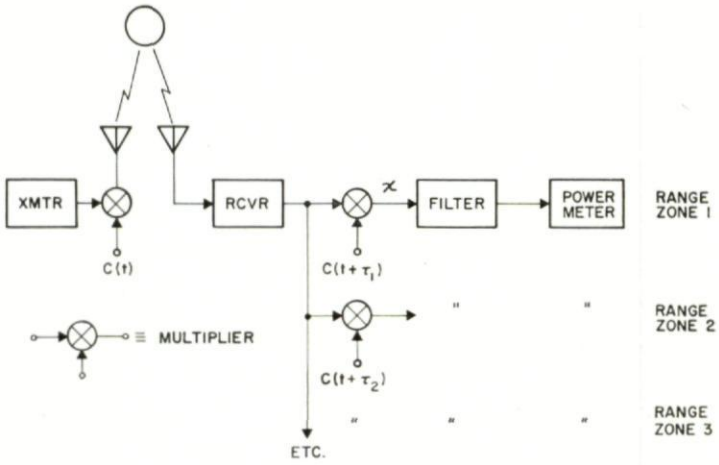


FIG. 8. *Range-gate block diagram.*

Figure 8 is a block diagram of this method. Amplitude modulation is considered here; but frequency, single sideband, etc., modulation may be and has been used successfully. The performance of this system depends critically on the waveform chosen for $C(t)$. The ability of the system to distinguish between different range zones (or time delays) depends on the autocorrelation function $R(\tau)$ of $C(t)$, defined by

$$R(\tau) = \text{average } [C(t)\, C(t+\tau)]$$

For a time-of-flight (range) resolution of τ_0, $R(\tau)$ must be a highly peaked function for $|\tau| < \tau_0$. To eliminate "cross-talk" between range zones, $R(\tau)$ should be near zero for $|\tau| > \tau_0$.

Two commonly used $C(t)$ and their autocorrelation functions are given in Fig. 9. The first waveform, $C_1(t)$, is a high, narrow pulse; the second, $C_2(t)$, is a random square wave, such that after each time interval of τ_0, a random choice is made as to whether $C_2(t) = \pm 1$ for the next interval. Both of these waveforms have the same autocorrelation function and the

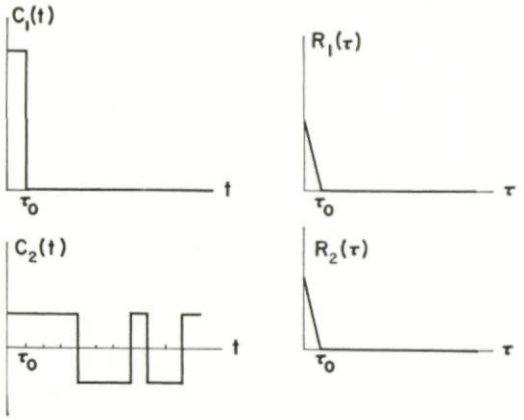


FIG. 9. *Modulation waveforms and their autocorrelation function.*

same range resolution, but they lead to entirely different radar hardware; $C_1(t)$ leads to a pulse radar, with its energy concentrated into narrow pulses, and $C_2(t)$ leads to a CW radar, with constant power over a cycle.

Consider for the following discussion that the target is a perfect mirror, so that the received signal is only shifted in time, and attenuated, from the transmitted signal. This restriction will be removed subsequently when we consider range-velocity mapping. The signal spectra at point x of Fig. 8 are given in Fig. 10a for the "correct" range zone, i.e., for the zone which contains the signal.

The pulse radar produces a line spectrum, so that the optimum filter for the power meter is a "comb" filter, which allows the lines to pass through, but blocks as much of the noise as possible. The spectrum for the CW radar is a single impulse. The multiplier at the receiver serves to cancel, exactly,

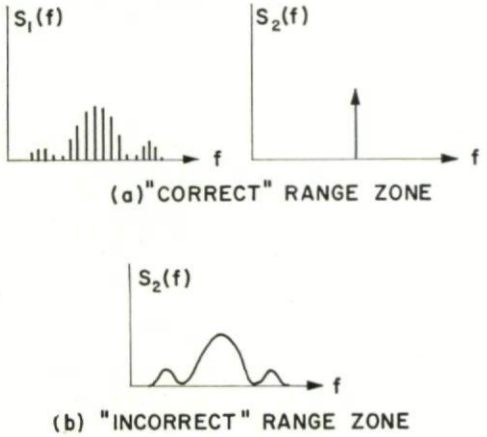


FIG. 10. *Range-gated spectra, point target.*

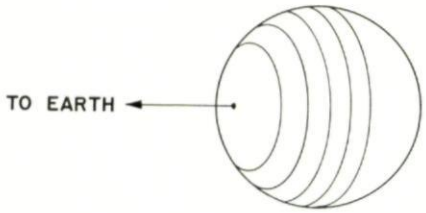


FIG. 11. *Contours of constant range.*

the modulation impressed upon the signal at the transmitter. Thus the signal is restored to a single sine wave. The optimum filter for this case is a narrow bandpass filter. There will be no signal spectrum in the "incorrect" range zones of the pulse radar, only noise. The corresponding spectrum for the CW case is given in Fig. 10b. It is still wide-band, so it is largely eliminated by the narrow-band filter. The remainder is, in practice, negligible when compared to the ever-present noise. It is interesting to note that for the same *average* power transmitted, $C_1(t)$ and $C_2(t)$ lead to radars of the same performance in estimating range.

Thus the set of range-gates divides the signal power according to its range composition. The contours of constant range are, of course, concentric circles centered about the sub-Earth point, as is illustrated in Fig. 11.

5. RANGE-VELOCITY MAPPING

The techniques of range-gating and spectral analysis may be combined simultaneously to produce a two-dimensional map of the surface of the target planet. The coordinates of such a map are along the contours of constant frequency shift and along the contours of constant range. Figure 12 illustrates this coordinate system. This system has an essential ambiguity, however, since there are, in general, two points on the planetary surface

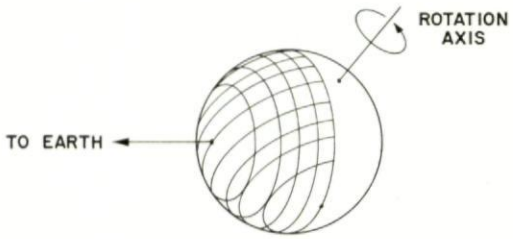


FIG. 12. *Contours of constant range and constant frequency shift.*

which have the same range and same line-of-sight velocity. This ambiguity may be resolved by observing the target planet at different times, when the rotation axis makes different angles to the line of sight.

The mechanization of this scheme, for the CW radar, is the same as for the range gates of Fig. 8, except that the power meters are replaced by spectrometers. A similar mechanization[10] can be used for the pulse-radar

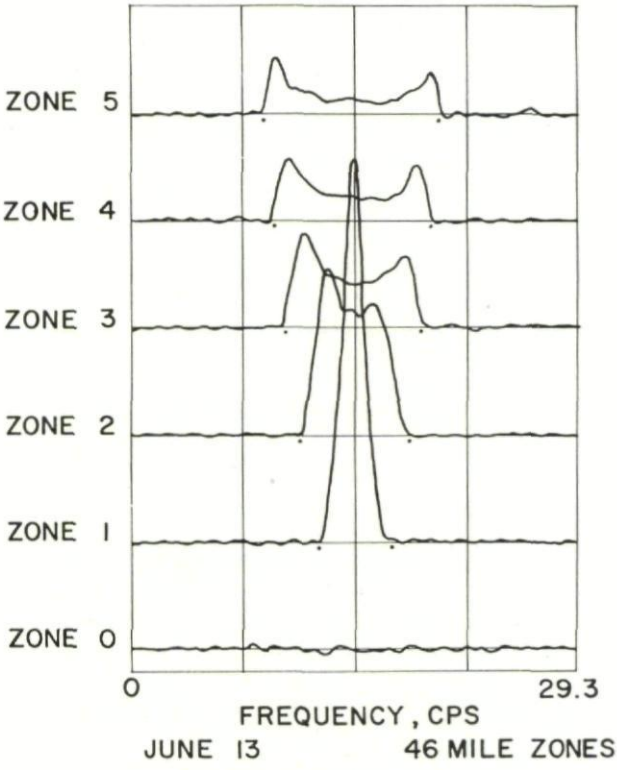


FIG. 13. *Range-gated Venus spectra, June 13, 1964, 46-mile zones.*

case. Each range-gate responds to only that echo which originates from the selected range zone and restores that signal to its sine wave form. The frequency of that sine wave depends, of course, on the line-of-sight velocity of the reflecting surface element. When this echo is resolved into its frequency spectrum, the two-dimensional mapping process is completed.

The theoretical limit of resolution in the range dimension and, simultaneously, in the frequency dimension is limited only by the stability of the oscillators of the system. In practice, the resolution attainable is limited by the signal-to-noise ratio. As the echo power is divided into smaller and smaller "cells," it finally cannot be distinguished from the background noise.

Figures 13 and 14 are actual samples of range-gated spectra of Venus echoes. Only the first few range zones near the sub-Earth point are shown. The spectra of all but the first few have the "double humped" shape characteristic of the coordinate system used and which stems from the oblique intersection of the coordinates near the frequency extremes. Any departures from this characteristic shape may be interpreted as features on the target's surface. As the planet rotates, permanent features will trace their signatures

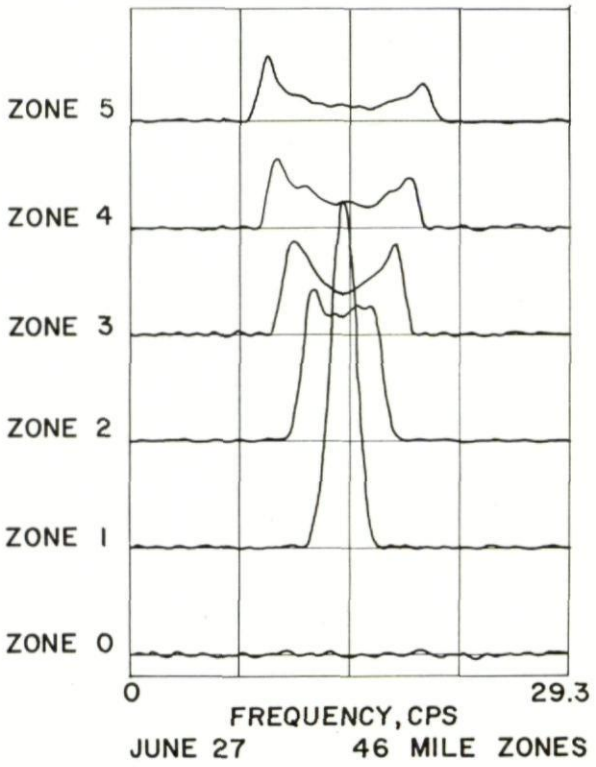


FIG. 14. *Range-gated Venus spectra, June 27, 1964, 46-mile zones.*

upon the spectra by moving from the high- to the low-frequency side, and, depending upon the feature's latitude, through the range zones.

Venusian features may be seen in Figs. 13 and 14. The peaks which appeared in zones 2 and 3 on June 13, 1964, had moved to become peaks in zones 4 and 5 by June 27, 1964.

The base bandwidths of the spectra (i.e., the frequency interval corresponding to the extreme frequency contours for the given range zone), measured over a time interval of several months, may be used to determine the rotation period of Venus and the direction in space of the axis. The bandwidths predicted from the JPL observations of 1962 (250 days retrograde) have been plotted on the figures.

The round-trip time-of-flight can be measured from these spectra with surprising precision by noting the positions and widths of the echoes in the range zones.

For example, at 24:00 GMT, June 13, the time-of-flight was 294.17890 $\pm$.00012 sec. Measurements made with 11-mile range-gates have yielded a precision of $\pm$.00003 sec.

DISCUSSION

————: I would be curious to know a little bit about the equipment that you used to make these measurements if it is possible to speak of it. Could you give two or three minutes on this, frequencies, hours so forth?

R. M. GOLDSTEIN: Each sector was taken at the Goldstone tracking station of JPL in the Mohave desert of California. The transmitter power was 100 kw average. The modulation of the range gear was a rapid switching of the phase of the CW carrier, the parabolic antenna is 85 ft in diameter, the wavelength is 12½ cm, the receiver noise temperature is about 30°K. The antenna gain both ways, allowing for all the losses, is about 108.5 db.

REFERENCES

1. DEWITT, J. H., JR., and STODOLA, E. K., Detection of radio signals reflected from the Moon, *Proc. Inst. Radio Engrs., N.Y.* **37,** 3, 229–242, 1949.
2. BAY, Z., *Hung. Acta Phys.*, **1,** 1, 1946.
3. VICTOR, W., STEVENS, R. and GOLOMB, S., Radar exploration of Venus, *Technical Report* **32–132,** Jet Propulsion Laboratory, Pasadena, 1961.
4. PETTENGILL, G. H., *et al.*, A Radar investigation of Venus, *Astron. J.*, **67,** 181–190, 1962.
5. THOMSON, J. H., *et al.*, A new determination of the solar parallax by means of radar echoes from Venus, *Nature, Lond.* **190,** 4775, 519, 1961.
6. KOTELNIKOV, V. A., Radar contact with Venus, *Brit. J. Inst. Radio Engrs.* **22,** 4, 293–295, 1961.
7. MUHLEMAN, D. O., *et al.*, The astronomical unit determined by Radar reflections from Venus, *Astron. J.*, **67,** 4, 1962.
8. GOLDSTEIN, R. M., and CARPENTER, R. L., Rotation of Venus: period estimated from radar measurements, *Science*, **139,** 3558, 910–911, 1963.
9. BLACKMAN, R. R. and TUKEY, J. W., *The measurement of power spectra*, Dover, New York, 1958.
10. PETTENGILL, G. H., Radio frequency scattering from the surface of the Moon, *Proc. Inst. Radio Engrs. N.Y.* **48,** 932–934, 1960.

Acknowledgment.—The author gratefully acknowledges the work of R. L. Carpenter and D. O. Muhleman, who have been his colleagues throughout the radar experiments, and H. C. Rumsey, who inverted the integral of Eq. (9).

CHAPTER 4

PULSE COMPRESSION RESEARCH IN THE UNITED KINGDOM

E. H. Boyenval
Royal Radar Establishment, Malvern, England

Experimental transmissions have been made with alternate pulses of single frequency and with frequency modulation. These have shown the marked advantage of even a crude pulse compression system over the normal radar when clutter returns are present. The design of an experimental pulse compression system is discussed. In this a balance was struck between the number of filter frequencies in the filter bank and the number of overlapping oscillator waveforms. A distortion measuring equipment has been developed which displays amplitude and phase errors measured on a single pulse. An adjustable compensating network has been developed.

The use of within-pulse frequency modulation to achieve pulse compression is now an accepted radar technique, and the various ways of achieving pulse compression have been widely described and discussed in many papers. The principal advantage of pulse compression is enhanced range resolution without the attendant engineering difficulties of producing pulses of very high peak power and very short duration. The principal disadvantages are associated with the production of spurious responses, or side lobes, and much time and effort has been spent in reducing and controlling these spurious responses, both in range and velocity.

As much of the original work was done in the United States, those engaged in pulse compression work in the United Kingdom have made no attempt to cover the whole field nor to investigate all the techniques open to the designer of pulse compression radars. Instead, efforts have been concentrated on a few points of particular interest.

Historically, the Bell Laboratories Chirp System was one of the first forms of pulse compression, and early work in the United Kingdom has followed on the lines pioneered by Bell Laboratories. The Marconi Company, sponsored by the Royal Radar Establishment, started a programme of research which included, among other items, the development of dispersive networks for use with Chirp systems. Mr. Brandon, in a later paper,* describes this work in detail, but this paper will discuss some early work on a large ground radar transmitter.

* Chapter 26

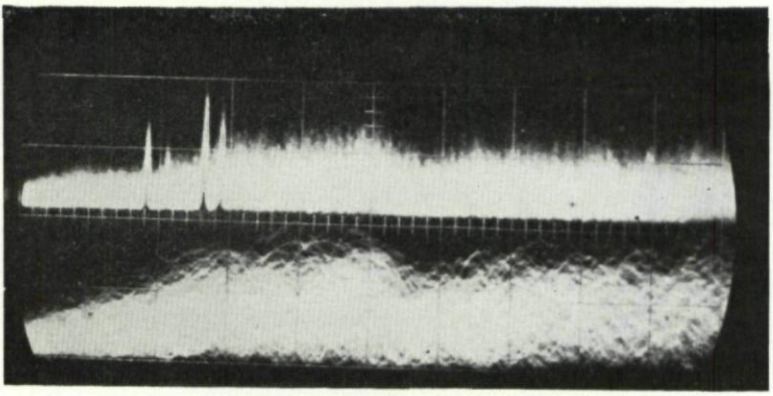
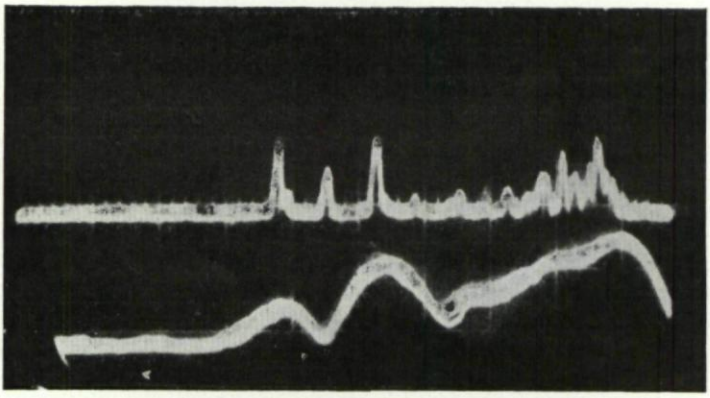
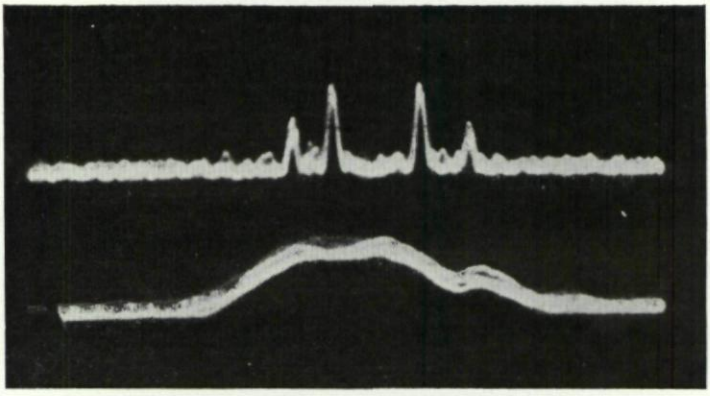


FIG. 1. *Performance against rain.*



(a) *Electricity pylons*



(b) *Rugby radio masts*

FIG. 2. *Performance against clutter.*

A very early model dispersive line was inserted in the I.F. stages of an existing radar transmitter. The linear frequency sweep was generated actively by applying a linear voltage ramp to the reflector of the master oscillator klystron.

Arrangements were made to radiate alternately a 5 μs single-frequency pulse, and a 5 μs pulse frequency modulated over a 5 Mc/s bandwidth. The I.F. and video bandwidths were increased appropriately, and the outputs using the two signals were displayed one above the other on a dual-trace A-scope display. Sidelobe levels were measured, and found to be -20 db.

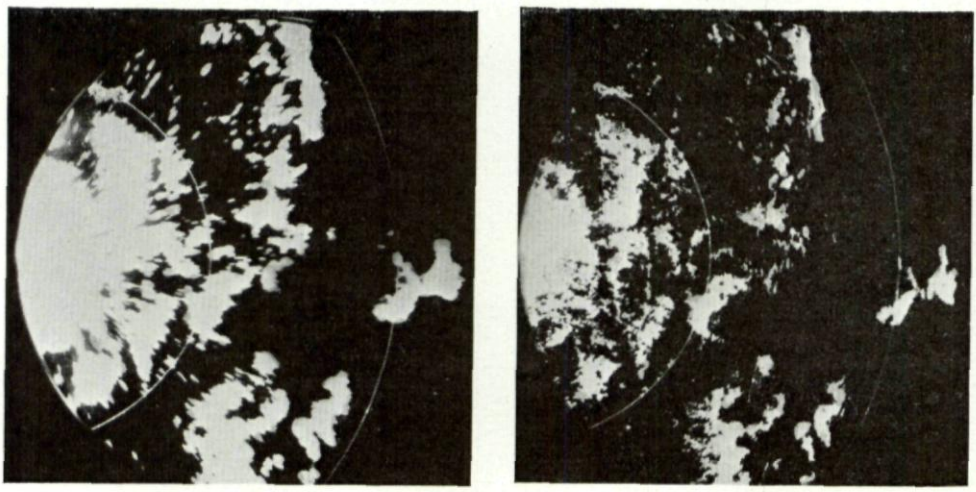


FIG. 3. *P.P.I. displays showing reduction of clutter.*

The reduction of the masking effect of rain clutter when using even this crude form of pulse compression is illustrated in Fig. 1, where the upper trace shows the radio masts at Rugby clearly visible despite the presence of rain echoes which completely mask the returns on the lower trace.

In Fig. 2 the increased resolution of small closely-spaced targets is shown; again, the pulse compressed signal is on the upper trace.

Figure 3 shows two p.p.i. pictures. The left hand picture was taken with the 5 μs single-frequency pulse, while on the pulse compressed picture on the right, the finer clutter detail is apparent. On this picture a power line and its supporting pylons can be clearly seen running across the top centre of the picture, and then continuing downwards across the centre.

Another field of activity at the Marconi Company, again under R.R.E. sponsorship, was the development of a correlation pulse compression system. Originally, the system was developed for use in a multichannel system, where it was thought that the problems of producing a number of dispersive networks of identical characteristics and all matched to the transmitted waveform would be formidable. A correlation system, making a direct

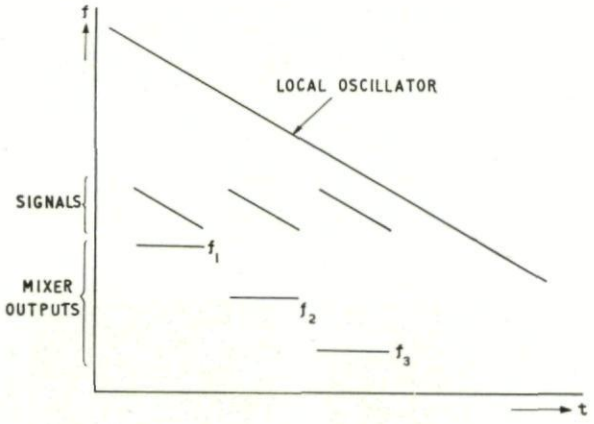


FIG. 4. *Single sweep local oscillator.*

comparison between the transmitted signal and the returning echoes avoided this difficulty, and one reference signal could be repeated in any number of correlators.

In the correlation pulse compression system, the transmitted signal is actively generated by a varactor-controlled frequency swept oscillator, sweeping from 60 Mc/s to 45 Mc/s in 10 μs. This sweep is repeated every 12.5 μs, and in a radar system, one sweep is gated out to modulate the transmitter.

If, on reception, the signals were mixed with a local oscillator sweeping continuously at 15 Mc/s/μs, a signal from a point target would give a constant frequency signal at the mixer output, lasting for 10 μs as shown in Fig. 4. The frequency of this signal would be a measure of range, and it would be possible to use a large bank of filters to give range resolution.

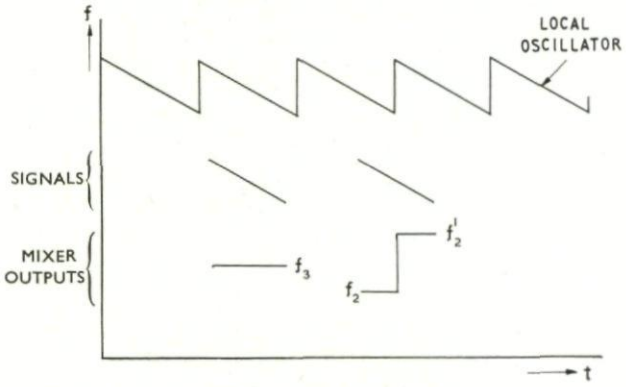


FIG. 5. *Repetitive local oscillator.*

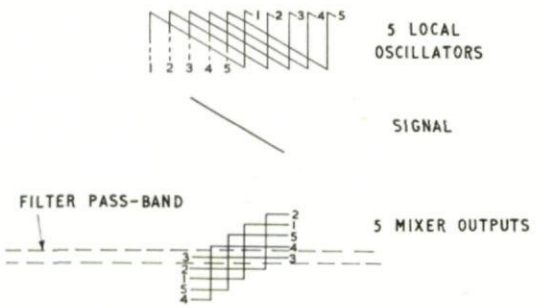


FIG. 6. *Multiple local oscillators.*

For a radar with a p.r.f. of 500 c/s, the local oscillator would have to sweep over a band of 3000 Mc/s. To achieve range discrimination commensurate with the transmitted bandwidth, filters would need to be 100 kc/s wide, and the filter bank would require 30,000 filters. Such a scheme is obviously impractical.

It is possible, however, to use the repetitive local oscillator waveform, as in Fig. 5. Target returns now either line up exactly with a local oscillator sweep or coincide partly with one sweep and partly with the next producing signals as shown in the figure. All the signal returns now lie in a much narrower frequency band, requiring less filters to cover it.

Further simplification follows if, instead of using one local oscillator waveform and a bank of filters all at different frequencies, we use a number of local oscillators, overlapping in time, and filters all at the same centre frequency. This is shown in a simplified form for 5 channels, in Figs. 6 and 7. The local oscillator waveforms are obtained from a tapped delay line, tapped at 2 μs intervals. At each tap the signal is mixed with the local oscillator, and the mixer output passed through a band-pass filter, which in this simplified case would have a bandwidth of 3 Mc/s.

Return signals whose front edge lay between $t = 0$ and $t = 2$ would produce a signal lying within the filter pass band when mixed with the first local oscillator, and would emerge from output A. Signals whose leading edge lay between $t = 2$ and $t = 4$ would emerge from output B, and those whose leading edge lay between $t = 4$, $t = 6$; $t = 6$, $t = 8$; and $t = 8$, $t = 10$ would emerge from outputs C, D and E. All these signals would, however be 10 μs in length, although the correlator has acheived resolution equivalent to a 2 μs pulse. To provide a conventional A-scope display, these outputs are all sampled in time-sequence by a 2 μs sampling pulse, derived in this simple example, from another delay line, and the outputs are all added on the output line. It is only necessary to provide sufficient equipment to cover one long pulse interval.

In the final system, 192 channels were provided. From the foregoing explanation, it follows that the number of different filter frequencies in the filter bank can be traded against the number of overlapping oscillator wave-forms. At one extreme, we use a single local oscillator and many

filters—as many as the final compression ratio. At the other extreme, we use many overlapping local oscillators, and one filter frequency. In the final design, a compromise solution resulted in the provision of a delay line with 64 taps. At each tap, the signal is mixed with the local oscillator, and then passed to 3 bandpass filters, centred on 19.9 Mc/s, 20.0 Mc/s and 20.1 Mc/s with 100 kc/s bandwidth.

The sampling pulses are effectively spaced at 65 ns, and are so chosen in shape and width that a target moving in range gives a smooth transition from one gate to the next. The system is shown in Fig. 8.

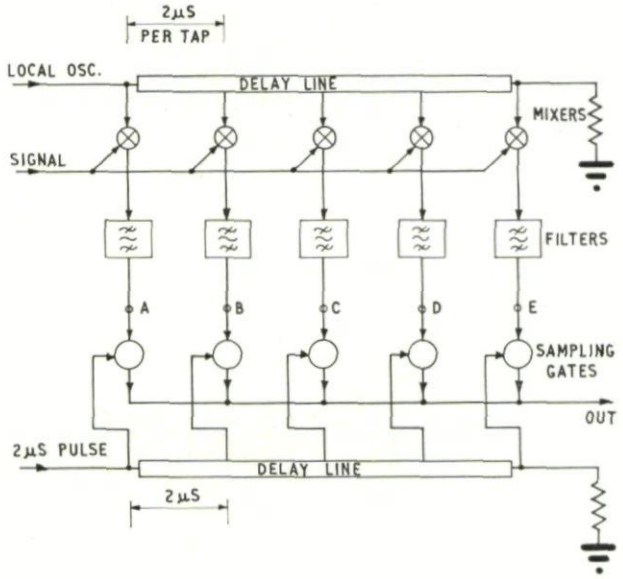


FIG. 7. *Fire channel pulse compression system.*

The complete system has been built at I.F. only, and demonstrated with simulated targets and noise.

The filter unit and mixer developed for the correlation pulse compression equipment has also been used to provide a range tracking system, which is applicable to any pulse compression radar using linear f.m. A single range tracker contains a local oscillator providing a single swept signal at a p.r.f. slightly different from the radar system p.r.f. This difference causes the range tracker to run through the system in range at a velocity of about 13,000 kts. On a target entering the tracker range interval, its appearance in the central (20 Mc/s) filter unit sets the tracker to the "track" mode. In this mode, the outer filters (19.9 Mc/s and 20.1 Mc/s) are used as a discriminator to control the p.r.f. of the oscillator, so that it tracks the target in range. When using more than one tracker, there is provision to lock out further trackers once a particular target has been acquired. The adjustment of tracker oscillator p.r.f. is initially achieved by a varactor

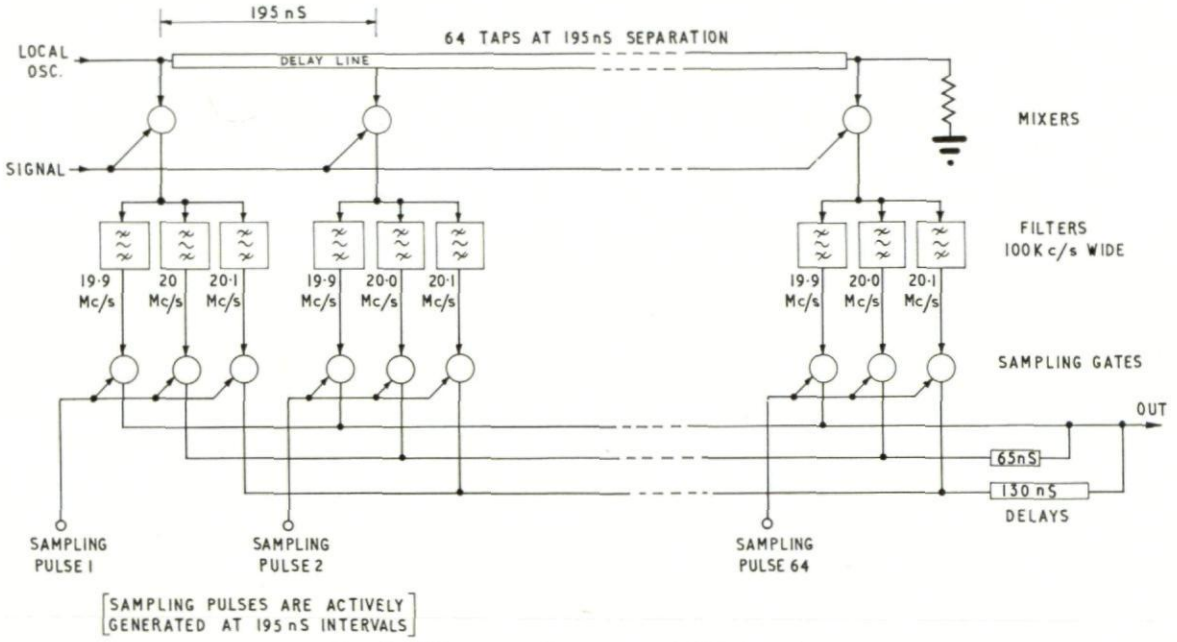


FIG. 8. *Pulse compression system with* 192 *channels.*

diode, but control is later taken over by a motor-driven capacitor, whose position then becomes a velocity memory for the particular target. Should the target become obscured by clutter, or fade, the tracker will follow in range at the same velocity for about 10 seconds, before returning to the search mode.

This equipment has again been built and demonstrated, at I.F. only, using simulated targets and noise. It has been shown to acquire satisfactorily targets 14 db below noise level moving with velocities of 2,000 kts., and two trackers have been shown tracking crossing targets whose relative velocity was as low as 100 kts., and with a difference in amplitude of 20 db without either tracker losing its target as they crossed over.

In any radar system using wide time-bandwidth product waveforms, there is no point in generating waveforms which approximate to the ideal for a particular form of pulse compression, and making decoding filters which are as close as possible to the matched filter for the chosen waveform if, between the function generator and the signal processor, the signal has to pass through devices which introduce frequency dependent phase and amplitude distortion. Both these forms of distortion are likely in travelling wave tubes and other high power devices, and the first step towards controlling the factors causing distortion or compensating for inevitable distortion is to have measuring equipment capable of assessing the distortion produced. Pulse to pulse variations make it very desirable to make such measurements on a single pulse, and the Marconi Company have developed measuring equipment to do this.

The original experimental model compares two signals, measuring relative phase with an accuracy of 0.4° and relative amplitude with an accuracy of 0.6%, for a signal sweeping through 20 Mc/s in 4 μs, with a centre frequency of 400 Mc/s. The measurement is completed in one pulse, and the equipment can handle error signals up to 100° in phase and 10 db in amplitude. This model is now working as a laboratory prototype. The results are presented on two C.R.T.'s, one for phase and the other for amplitude, as traces superimposed on a calibration graticule. The horizontal coordinate in each case is frequency, and the vertical coordinate phase angle or amplitude in db. The graticule lines provide calibration markers, which are derived through the same circuits as those used for measurement. The equipment normally operates in the calibration mode, and is switched to the measurement mode by a pre-pulse preceding the signal.

The basic circuits for the measurement of phase and amplitude are shown in Figs. 9 and 10.

All the measurement is done at intermediate frequency, and a much more versatile instrument is now being developed from the first model. It will have a number of interchangeable R.F. heads and will cover the frequency range 2000 Mc/s to 10,000 Mc/s in several bands. It will cope with a frequency sweep of up to 1% of the centre frequency, and with sweep rates of up to 10 Mc/s/μs, provided that the time-bandwidth product does not exceed 2000. Sweep duration may vary from 2 μs to 5 ms. With this equipment, it should be possible to assess accurately the distortions arising in pulse compression radar systems.

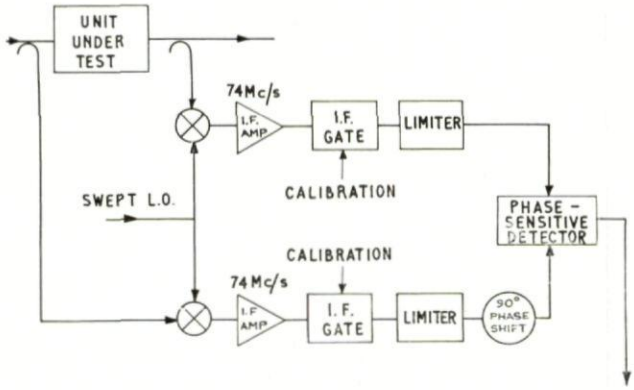


FIG. 9. *Phase measurement.*

To provide some compensation for unwanted sidelobes arising from distortions introduced by components in a radar system, an adjustable compensating network has been developed. It is shown diagrammatically in Fig. 11. The main signal is injected at the centre point of a delay line. By combining in-phase and quadrature versions of the signal, a replica of the signal, whose phase can be varied through 360° and whose amplitude is under control, can be added at various taps on the same line, either in advance of, or delayed behind, the main signal. The time spacing of the taps is at intervals of 1/B, where B is the signal bandwidth, and the amplitude of the correction which can be applied is limited, the total available correction decreasing away from the centre tap.

As Mr. Brandon will show in his paper, the design of lumped constant lines for use in linear f.m. systems with bandwidths of some 15 to 20 Mc/s is now on a firm basis. Future work will be directed towards the production of dispersers with bandwidths up to 200 Mc/s, but work in this field has not yet passed beyond the stage of preliminary assessment.

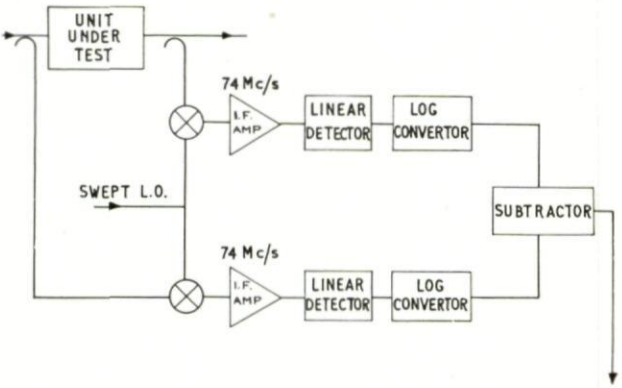


FIG. 10. *Amplitude measurement.*

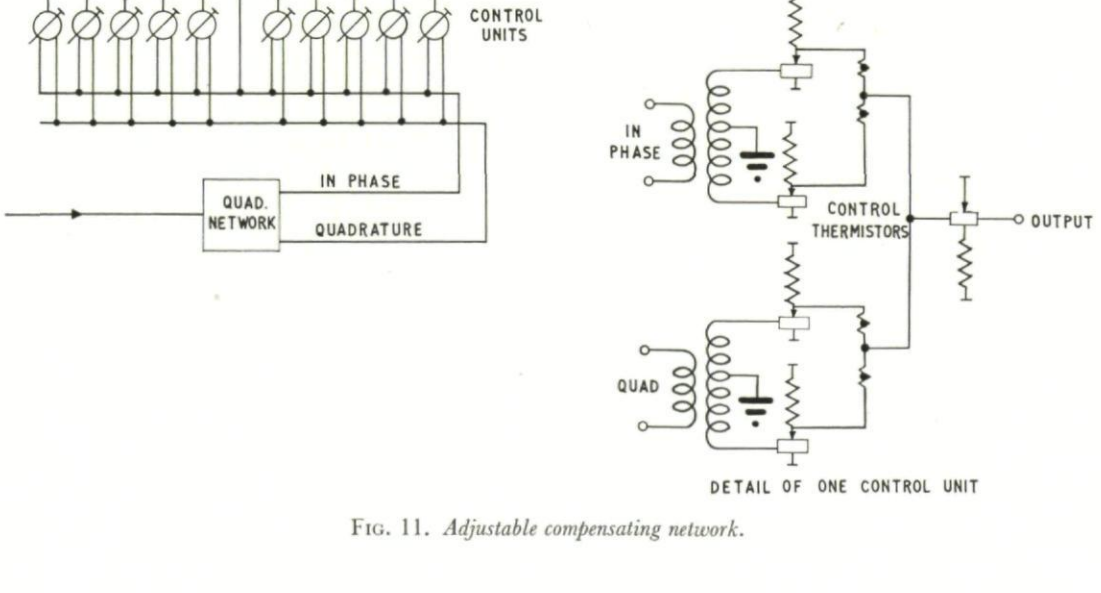


FIG. 11. *Adjustable compensating network.*

DISCUSSION

R. J. LEES: These radars that you are talking about are getting very complicated with lots and lots of components, are they going to be reliable, do you think?

E. H. BOYENVAL: I think on the linear FM chirp type of radar the majority of the decoding components are passive, such as the bridged-T type of circuit which once adjusted and set are pretty stable. I don't think one is introducing a great deal of what one might call "failure prone equipment" into the radar by adopting this kind of technique.

E. F. DUKES: I will speak for the non-professional here as I am not a radar man but the first speaker to-day commented on the fact that the target changed the form of the signal. In your paper you are sending out a signal of a very special form and operating with it when it returns. Could you comment upon the effect of the target on the interpretation of pulse compression returns?

E. H. BOYENVAL: I think it is certainly true that any target characteristics which are going to introduce variation in the phase relations between components of the reflection from the target are of course going to deteriorate the efficiency of such a pulse compression. How much this will have an effect I don't know. I think it more likely to arise in producing an increase in the unwanted side lobe responses. Of course this raises the whole question of what happens to pulse compression systems in the presence of Doppler shifts. In many linear FM systems there is a marked coupling between velocity and range and a Doppler shift on the return signal is converted directly to range error.

Y. FOUCHÉ: Vous avez indiqué qu'un appareil de mesure de phase et d'amplitude tres précis avait été mis au point pour les émetteurs destinés à la compression d'impulsion.

(i) Pourquoi cette précision est elle nécessaire?

(ii) Cet appariel peut il servir à mesurer les caractéristiques des récepteurs et en particulier des lignes dispersives?

E. H. BOYENVAL: I don't think it matters very much whether frequency dependent phase shift gets on to this type of signal either in the transmitter or the receiver, the effect is very much the same. You are attempting to transmit a signal which has a particular law of phase against time, or frequency against time, which ever you wish, and you are receiving this in a filter which is pre-matched to such a signal. Now if you do anything to it anywhere between the place where you generate the signal and the place where you process it which alters this relationship, the effects are going to be less than optimum, you are not going to achieve the results which you hope. The equipment we are developing was designed primarily to look at distortions occurring in high power transmitters, but there is no reason why it should not operate on distortion occurring in devices in the receiving channel as well. The second point of your question was whether or not it is necessary

to measure this phase error. The phase has to be controlled to within certainly one or two degrees over the band if one does not want to produce quite large unwanted sidelobes. Compensating networks can be relied upon to remove some of the unwanted effects but it is a policy of dispair to say 'never mind about the distortion occurring you can always take it out on an adjusting network afterwards.' I think one wants to measure the transmitted signals accurately first in order to see just what sort of problem one is up against.

CHAPTER 5

METHODES STATISTIQUES EN RADIODETECTION

L. REBOUL
C.N.E.T., Issy-les-Moulineaux (Seine), France

A theory of detection is presented which leads to a criterion for optimum detection, that is the highest probability of detection for a given false alarm rate. In this analysis the problem is divided into four parts of increasing complexity:

(1) *The ideal problem of detection. In this case there is the possibility of a single known target and the decision is whether the target is present or not, i.e. whether we have a known signal plus noise or noise alone.*

(2) *The stochastic problem of detection. This is an extension of the previous case. Any one of a number of targets may be present. There is no knowledge of what is sought, the decision is between "targets" and "no targets."*

(3) *The problem of estimation. This extends* (2) *to the determination of which particular target is present and to the study of its characteristics.*

(4) *The problem of discrimination and resolution. This covers the resolution of one target from another and of a target from noise and clutter.*

Nos connaissances actuelles en logique, statistique, théorie des décisions et bien d'autres éléments importants ne permettent pas d'élaborer de théorie cohérente englobant tous les aspects du problème, de systèmes complexes tels que les radiodétecteurs. Nous avons à notre disposition une collection de fragments orientés vers des problèmes particuliers. On effectue les projets à l'aide d'études paramétriques et notre connaissance des facteurs qui entrent en jeu ne nous permet pas de formuler de définition de la meilleure méthode de synthèse.

Mais ces connaissances, en statistique et en théorie des décisions, nous les avons. Nous ne sommes plus réduits à l'equation du radar, qui était le seul outil mathématique disponible à la fin de la deuxième guerre mondiale, cette équation qui permet seulement le calcul de la sensibilité du système pour une impulsion unique, et qui est donc parfaitement inadéquate même pour la synthèse de dispositifs de tracking sur cible unique.

METHODES STATISTIQUES DE SYNTHESE DES RADIODETECTEURS

La grande majorité des radars envoient des impulsions dans l'espace et reçoivent des échos. Le but est de déterminer la réponse percussionnelle de ce réseau qu'est l'espace avec ses différentes sortes de cibles et d'anomalies.

Le problème est donc pour chaque catégorie de réseau à étudier, de choisir au mieux le signal à émettre et les opérations à lui faire subir à la réception pour avoir le plus d'information possible sur le dit réseau. La synthèse d'un radar comprend donc le choix du signal émis et la conception de la partie du récepteur destinée à manipuler l'information reçue.

Evidemment la solution générale de ce problème est quasi impossible et nous pouvons seulement analyser l'effet de certains signaux et non pas décrire une méthode pour choisir le meilleur signal possible pour une application donnée.

Nous parlerons des problèmes suivants:

(1) *Le problème idéal de la détection* correspond à une situation extrêmement idéalisée dans laquelle une cible de caractéristiques connues est supposée, soit présente, soit absente. On doit recevoir, soit un signal connu et du bruit, soit du bruit seul.

(2) *Le problème stochastique de la détection* qui est une extension du problème précédent aux cas où l'*une* quelconque, parmi un certain nombre de cibles connues, peut être présente (ou le cas où la cible a des paramètres aléatoires) mais sans s'occuper de préciser quelle cible particulière ou la valeur du paramètre. L'information désirée reste simple: "des cibles" ou "pas de cibles."

(3) *Le problème de l'estimation.* Extension du cas précédent en cherchant à préciser quelle cible est présente ou quelle est la valeur du paramètre.

(4) *Le problème de la discrimination et du pouvoir séparateur* qui consiste à reconsidérer les problèmes précédents dans les cas où le "bruit" est constitué en partie par des échos en provenance d'autres cibles et par suite, dépend du signal émis.

1. LE PROBLEME IDEAL DE LA DETECTION

Le problème le plus simple pour illustrer ces méthodes mathématiques, leurs avantages et leurs difficultés, est peut-être le suivant: supposons que nous sachions que l'espace qui entoure le radar puisse prendre deux états :

Etre vide ou contenir une seule cible.

Nous supposons, en outre, que les caractéristiques de cette cible (si elle est présente) surface effective, distance, azimuth, vitesse, etc . . . , sont connues, ainsi que tous les paramètres de l'émetteur, de telle sorte que si la cible est présente, nous pouvons, en ignorant le bruit, calculer le signal $s(t)$ qui apparaîtrait sur l'antenne de réception. Enfin, nous supposons qu'il y a à ces mêmes bornes d'entrée, un signal supplémentaire de bruit $n(t)$ correspondant à la température effective de bruit du récepteur, possédant une densité de puissance constante en fonction de la fréquence sur toute la bande du signal $s(t)$, et d'amplitude N_0 watts/hertz. Le signal effectivement observé est donc:

$$f(t) = s(t) + n(t) \quad \text{si la cible est présente}$$

ou $$f(t) = n(t) \quad \text{s'il n'y a pas de cible}$$

Le but du dispositif de manipulation des données est donc de déduire de l'examen de $f(t)$ laquelle des deux conditions a le plus de chance d'être vraie, (le plus de chance sera à préciser). La situation est indiquée schématiquement par la figure 1 qui suggère quelques remarques. Tout d'abord, nous supposons que le signal $s(t)$ est de durée *finie*. Ce qui traduit à la fois le fait que dans la réalité, les cibles apparaissent et disparaissent ou que le faisceau de l'antenne les passe, et le désir d'éviter le résultat, à peu près évident qu'avec un signal de durée infinie, et étant donné les hypothèses faites, on aurait peu de difficulté à faire des décisions parfaites. Le deuxième point est plus délicat: le système de manipulation des données prend des décisions et ne se contente pas de filtrer le signal reçu pour éliminer le plus de bruit possible. Sur la figure 1 la décision est suggérée sous la forme to be or not to be provenant du problème choisi. Une fois la décision faite, la

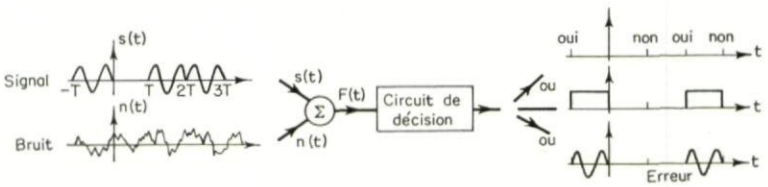


FIG. 1

façon de la présenter n'a aucune importance. Nous sommes très loin des présentations familières sur scope *A* ou *PPI* et la différence fondamentale d'attitude demande un examen plus détaillé.

Pourquoi le dispositif de manipulation des données est-il nécessaire? Pourquoi ne pas présenter le signal reçu tel que, à l'observateur, en lui laissant le soin des déductions? Il y a au moins deux raisons évidentes:

—l'une est une simple question d'adaptation. Les sens humains ne sont pas adaptés à l'observation de signaux électriques en bande *S;* c'est pourquoi nous les convertissons en signaux lumineux en video. En principe, cette transformation ne nous concerne pas, car elle n'apporte aucune perte d'information.

—l'autre est également une question d'adaptation: dans tout système radiodétecteur la vitesse (rate) avec laquelle l'information arrive sur l'antenne (j'utilise ici le mot information dans un sens non technique, pour désigner la vitesse à laquelle $f(t)$ prend des valeurs suffisamment distinctes) dépasse très largement la capacité de récepteurs, tels que calculateurs ou êtres humains. Par suite, un dispositif de manipulation des données est nécessaire précisément pour détruire l'information inutile dans le signal reçu sans perdre d'information utile. Opération essentiellement irréversible. C'est cette forme de manipulation des données qui nous intéresse. En terme de communications, dans un récepteur superhétérodyne nous nous intéressons non pas au premier et deuxième détecteur, qui ne modifient pas la quantité d'information mais plutôt aux caractéristiques de filtrage des amplificateurs *RF* et *IF* qui, par des

opérations irréversibles, éliminent du signal reçu l'information inutile (bruit et interférence).

Moyenne du Carré de la Différence

Comme premier exemple considérons une méthode arbitraire, mais apparemment raisonnable pour obtenir une décision du type de la figure 1. Il est logique de comparer le signal reçu $f(t)$ pour chaque intervalle de durée T, à $s(t)$. Si $f(t)$ "ressemble" plus à $s(t)$ qu'à zéro nous annoncerons "signal présent."

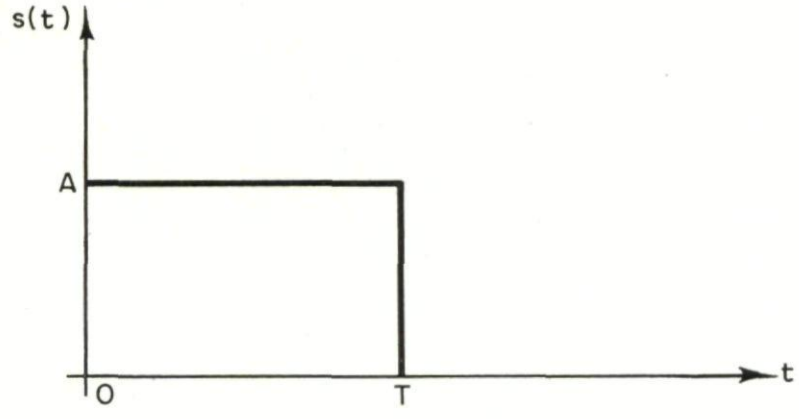


FIG. 2

Comme mesure de cette "ressemblance" prenons l'intégrale du carré de la différence, c'est-à-dire comparons sur l'intervalle $0 < t < T$:

$$\int_0^T [f(t) - s(t)]^2 \, dt \text{ à } \int_0^T [f(t) - 0]^2 \, dt$$

La règle sera donc:

$$\int_0^T [f(t) - s(t)]^2 \, dt \underset{OUI}{\overset{NON}{\gtrless}} \int_0^T [f(t)]^2 \, dt$$

Si je pose:

$$E = \int_0^T s^2(t) \, dt = \text{Energie contenue dans le signal (niveau d'impédance: un ohm)}$$

il vient en développant, la régle de décision équivalente:

$$\int_0^T f(t) \, s(t) \, dt \underset{NON}{\overset{OUI}{\gtrless}} \frac{E}{2}$$

La figure 2 montre comment peut être réalisée cette règle de décision.

Le filtre passe-bas effectue approximativement l'intégration. Si l'on fait passer $s(t)$, $f(t)$ ou les deux, à travers un circuit de déclenchement de telle façon que le produit $f(t)\ s(t)$ soit nul sauf pour $0 < t < T$, il suffit d'un filtre RC avec $RC \gg T$ pour effectuer approximativement l'intégration.

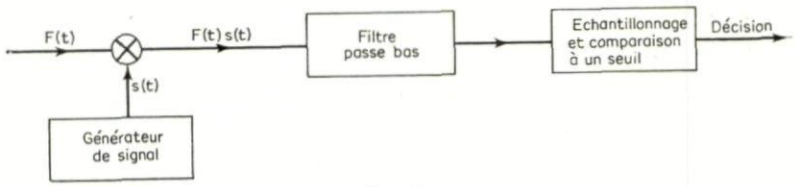


FIG. 3

Bien entendu si l'on veut examiner des intervalles successifs, il sera nécessaire d'avoir un autre filtre RC à mettre en route à $t = T$ pour l'intervalle $T < t < 2T$.

La commutation des filtres est un artifice assez lourd et il est souvent plus commode d'utiliser un filtre ayant une répose percussionnelle de la forme indiquée figure 3.

En fait, il n'est pas nécessaire dans la plupart des cas, que la réponse percussionnelle de la figure 3 soit parfaitement réalisée. Un simple filtre RC avec $RC \sim T$ est souvent satisfaisant.

Un dispositif tel que celui de la figure 2 est souvent appelé détecteur à corrélation parce que l'intégrale qu'il calcule est une mesure de la fonction de corrélation entre le signal reçu $f(t)$ et le signal attendu $s(t)$.

Les détecteurs à corrélation ont certaines propriétés optima que la discussion n'a certainement pas fait ressortir, car nous n'avons donné aucun moyen de comparer les règles de décision entre elles.

On peut imaginer d'autres règles de décisions et si l'on veut pouvoir les comparer théoriquement il faut faire des hypothèses sur la nature du bruit:

Densité spectrale connue et constante

Cette donnée ne définit pas complètement le bruit mais permet d'effectuer des comparaisons entre différents systèmes en terme du rapport S/B. Le choix des systèmes de détection et du rapport S/B comme critère de sélection reste arbitraire.

Nous allons restreindre notre discussion à des critères de décision qui peuvent être réalisés comme indiqué dans la figure 4.

Le filtre étant linéaire nous pouvons écrire:

$$g(T) = g_s(T) + g_n(T)$$

$g_s(T)$ et $g_n(T)$ sont les réponses à l'instant T du filtre aux signaux $s(t)$ et $n(t)$ Nous définissons le rapport signal sur bruit a, par:

$$a = \frac{g_s^2(t)}{\langle g_n^2(t) \rangle}$$

et nous cherchons la réponse percussionnelle $h(t)$ qui rend a maximum.

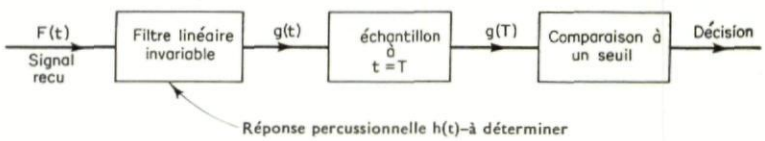


FIG. 4

On trouve $h(t) = s(T - t)$ ou $H(\omega) = S^*(\omega)e^{-j\omega T}$ où:

$$S(\omega) = \int_0^T s(t)\ e^{-j\omega T}\ dt$$

Un tel filtre est appelé le "filtre adapté" à $s(t)$. Avec le filtre l'opération de la figure 4 se résume à:

$$g(T) = \int_0^T f(t)h(T - t)\ dt = \int_0^T f(t)s(t)\ dt \underset{NON}{\overset{OUI}{\gtrless}} \text{seuil}$$

Ce qui montre que le filtre adapté et la détection par corrélation conduisent aux mêmes opérations mathématiques sur le signal et sont donc équivalentes.

La valeur maximum du rapport S/B est:

$$a = E/N_0 = \frac{\text{Energie du signal}}{\text{Puissance de bruit par hertz}}$$

Ce résultat est très important car il montre que les performances du système, dans la mesure où on les évalue par le rapport signal sur bruit, sont indépendantes des qualités du signal $s(t)$ autres que son énergie, et il y a bien longtemps que l'on a cessé de mettre en doute les qualités du rapport S/B comme critère de performance.

La discussion actuelle est plus intéressante parce que moins abstraite. Nous avons pu montrer que le filtre adapte ou la détection par corrélation a un caractère optimum. Mais il est important de se rendre compte que le pas en avant que nous venons de faire a été payé par l'hypothèse de la connaissance du spectre de bruit.

Pour éliminer l'arbitraire du choix du rapport S/B comme critère de performances il est nécessaire de préciser les hypothèses et de supposer posséder une description statistique complète du signal reçu, ce qui veut dire en gros que nous pouvons chiffrer la probabilité d'apparition de chaque forme particulière de tension de bruit $n(t)$.

Notre connaissance du signal reçu est correctement représentée par des échantillons suffisamment rapprochés. Si nous représentons les échantillons de $f(t)$ aux instants $t_K = k\ \Delta t$ par le vecteur:

$$\begin{aligned} F &= [f(t_1), f(t_2), f(T)] \quad \text{et} \\ N &= [n(t_1), \qquad n(T)] \\ S &= [s(t_1), \qquad s(T)] \end{aligned}$$

nous pouvons introduire diverses probabilités concernant le signal et le bruit, en particulier:

La probabilité a posteriori est que le signal soit présent en supposant avoir reçu un signal $f(t)$ particulier, $P[s + n/F]$.

On définit le rapport de vraisemblance Λ:

$$\Lambda = \frac{p[F/s + n]}{p[F/n]}$$

La quantité précédente $p[s + n/F]$ est une fonction monotone croissante de Λ. Or, d'après Woodward, la valeur de cette probabilité a posteriori est précisément le maximum de connaissance que l'on puisse avoir sur la situation réelle. Or, puisqu'elle est monotone en Λ, on peut dire que Λ résume en un seul chiffre toute l'information contenue dans $f(t)$ nécessaire à une décision sur la présence ou l'absence de $s(t)$. Passer du grand nombre de chiffres nécessaires à la spécification d'un signal $f(t)$ particulier, à un seul Λ sans perte d'information sur la présence de $s(t)$ est un grand pas dans la réduction des données. Mais ce n'est pas encore une décision. Il est clair toutefois qu'elle sera de la forme:

$$\Lambda \underset{NON}{\overset{OUI}{\gtrless}} \text{seuil } \gamma \text{ à préciser.}$$

Considérons un cas particulier, très important pour les applications et l'un des rares pour lequel on peut calculer Λ en fonction de F. Supposons que le bruit soit un processus gaussien stationnaire à densité spectrale constante N_0 watt/hertz, on montre que Λ s'écrit:

$$\Lambda = \exp\left[\frac{1}{N_0}\int_0^T f(t)\, s(t)\, dt\right] \exp\left[-\frac{1}{2N_0}\int_0^T s^2(t)\, dt\right]$$

En prenant les logarithmes et en absorbant les constantes dans le seuil, déjà arbitraire, on obtient la règle de décision:

$$\int_0^T f(t)\, s(t)\, dt \underset{NON}{\overset{OUI}{\gtrless}} \text{seuil}$$

C'est précisément la règle de décision obtenue avec le filtre adapté ou le corrélateur, mais justifiée maintenant comme conséquence du bruit blanc gaussien, comme règle de décision optimum sans autre caractère arbitraire que la valeur du seuil.

2. METHODES DE LA THEORIE DES DECISIONS

Pour supprimer ce degré final d'arbitraire, il faut préciser le but et pouvoir évaluer les performances du système de décision. On peut évidemment réduire le nombre d'erreurs. Mais ce n'est pas assez précis car notre système peut faire deux sortes d'erreurs:

—fausses alarmes
—cibles manquées.

Pour chaque système de décision et classe de signaux reçus on peut associer une probabilité à chacune d'elles:

P_F = probabilité de fausse alarme
P_M = probabilité de manquer la cible
ou souvent P_D = probabilité de détection
$= 1 - P_M$

On voudrait évidemment réduire ces deux quantités le plus possible, mais naturellement une réduction de l'une s'obtient au prix d'une augmentation de l'autre. Il y a donc un compromis à trouver.

Posé de cette façon, notre problème relève du domaine des tests d'hypothèse ou de la théorie de décision et fait l'objet d'une vaste littérature.

Le processus le plus ancien et le plus familier pour équilibrer les probabilités d'erreur est le critère de Neyman-Pearson. Ce critère propose, appliqué au cas présent, de fixer une des probabilités d'erreur, soit P_F à une valeur acceptable et de chercher la règle de décision qui rend l'autre minimum, ce qui conduit à une règle de décision:

$$\Lambda \gtrless \text{seuil } \gamma \text{ à choisir correspondant à } P_f$$

Un autre exemple de procédé est la solution de Bayes. Nous postulons que chaque signal manqué coûte C_M et que chaque fausse alarme coûte C_F. Alors le coût moyen de toute règle de décision sera:

$$C = C_M P_M P_{S+n} + C_F P_F P_n$$

Nous cherchons la règle de décision qui rend C minimum. Il est facile de montrer qu'elle sera de la forme:

$$\Lambda \gtrless \gamma = \frac{C_F P_n}{C_M P_{S+n}}$$

Performances et Sensibilité du Filtre Adapté

On peut facilement calculer les probabilités de détection et de fausse alarme pour la règle de décision qui utilise le filtre adapté pour toute valeur du seuil γ. Nous voulons calculer la probabilité que:

$$L = \int_0^T f(t)s(t)\,dt > \gamma \text{ dans les deux cas}$$

$$f(t) = n(t) \text{ et}$$

$$f(t) = s(t) + n(t)$$

où $n(t)$ est du bruit blanc gaussien. Puisque L peut être considéré comme la sortie d'un filtre linéaire de réponse percussionnelle:

$$h(t) = S(T - t)$$

L sera une variable aléatoire gaussienne dans les deux cas. Supposons que:

$$f(t) = n(t)$$

$$E(L) = 0$$

$$E(L^2) = \frac{N_0}{2\pi} \int_{-\infty}^{+\infty} |H(\omega)|^2\,d\omega = N_0 \int_0^T s^2(T - t)\,dt = N_0 E$$

où $E[\ ]$ est l'espérance mathématique.

Si $f(t) = n(t) + s(t)$ on trouve

$$E(L) = \int_0^T s^2(t)\,dt = E$$

$$E(L^2) = N_0 E + E^2$$

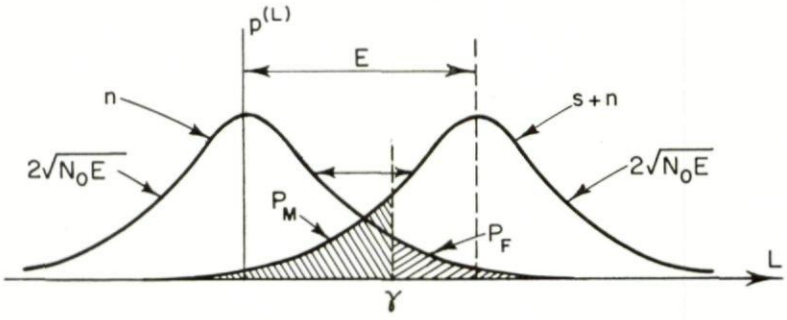


FIG. 5

La densité de probabilité de L a donc la forme indiquée par la figure 5.
Les aires des régions hachurées correspondent à P_M et P_f.
Analytiquement:

$$\begin{cases} P_f = \displaystyle\int_\gamma^\infty \frac{1}{\sqrt{2\pi}\sqrt{N_0E}}\, e^{-x^2/2N_0E}\, dx = \int_{\gamma/\sqrt{N_0E}}^\infty \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2}\, dx \\ P_D = \displaystyle\int_\gamma^\infty \frac{1}{\sqrt{2\pi}\sqrt{N_0E}}\, e^{(x-E)^2/(2N_0E)}\, dx = \int_{(\gamma-E)/\sqrt{N_0E}}^\infty \frac{1}{2\pi}\, e^{-x^2/2}\, dx \end{cases}$$

La figure 6 donne la loi de variation de P_f en fonction de γ.

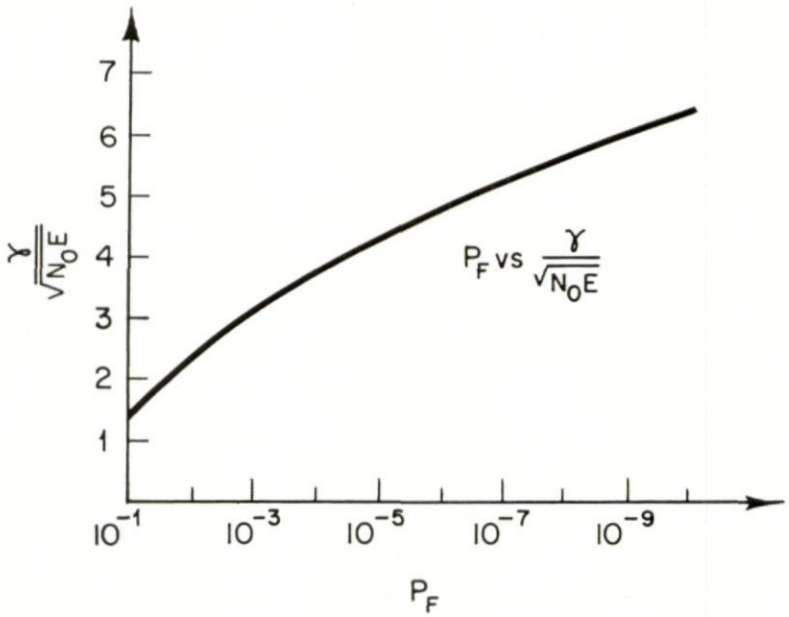


FIG. 6

On remarque que P_f est extrêmement sensible à de faibles variations de γ. C'est une des raisons principales de l'existence des dispositifs à taux de fausse alarme constant.

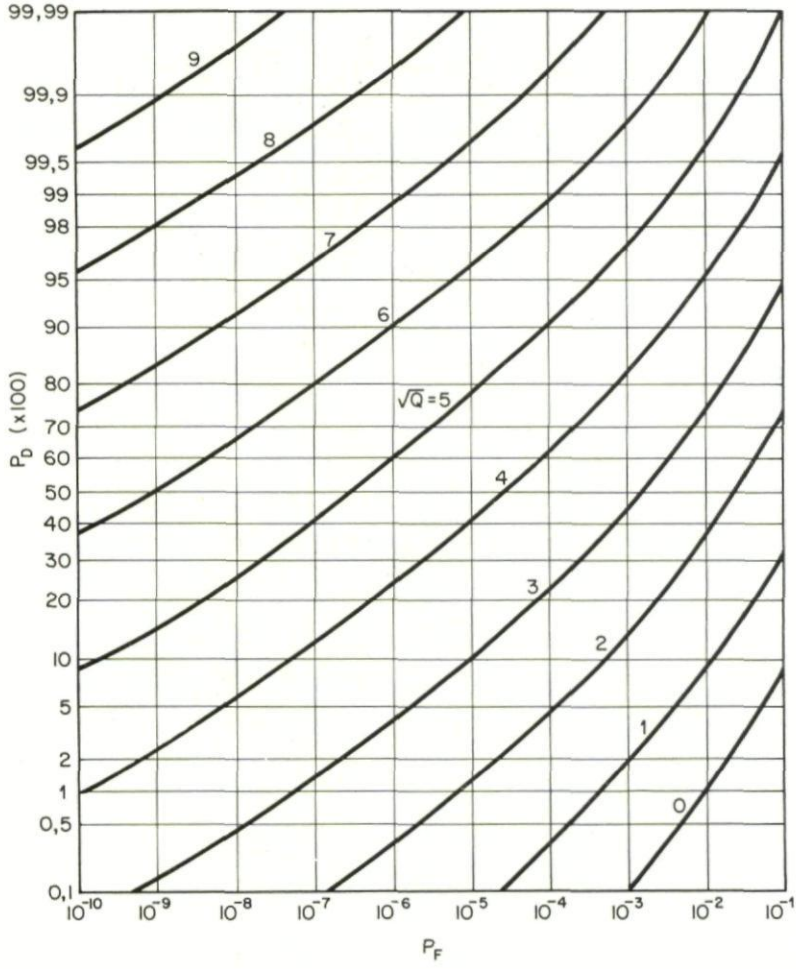


FIG. 7

On peut éliminer γ entre les deux relations qui donne P_f et P_D. On obtient un réseau de courbes pour différentes valeurs du coefficient $Q = \dfrac{E}{N_0}$ (Fig. 7).

D'après ces courbes P_D est assez insensible à P_f et Q varie lentement en fonction de P_D.

3. LE PROBLEME STOCHASTIQUE DE LA DETECTION

Aucun problème réel ne peut se réduire au modèle du paragraphe précédent. Nous supposons toujours que notre seul but est la détection c'est-à-dire présence ou absence. Il y a toujours une seule cible. Le signal peut être représenté par une fonction du temps et d'un nombre fini d'autres paramètres (amplitude, retard, fréquence, etc . . . .) qui sont fixés pour chaque intervalle de détection mais varient d'un intervalle à l'autre et seront considérés comme variables aléatoires:

$$f(t) = s(t, a) + n(t) \text{ ou}$$
$$f(t) = n(t) \text{ bruit blanc gaussien}$$

Le but du dispositif de manipulation des données est de décider si $s(t, a)$ est présent avec n'importe quelle valeur de a, ou non.

On démontre que la quantité $\int_{-\infty}^{\infty} \Lambda(a)\, p(a)\, da$ représente une statistique suffisante et que la règle de décision doit être de la forme:

$$\int_{-\infty}^{\infty} \Lambda(a)\, p(a)\, da \gtrless \gamma$$

Dans le cas du bruit blanc gaussien on peut calculer $\Lambda(a)$ comme dans le paragraphe précédent:

$$\Lambda(a) = e^{1/N_0} \int_0^T f(t)s(t, a)\, dt\, e^{(-1/2N_0)} \int_0^T s^2(t, a)\, dt$$

il vient donc:

$$\int_{-\infty}^{\infty} pa \exp\left[1/N_0 \int_0^T f(t)s(t, a)\, dt\right] \exp\left[-\frac{1}{2N_0} \int_0^T s^2(t, a)\, dt\right] da \gtrless \gamma$$

Le premier membre de cette relation peut être considéré comme la sortie à l'instant $t = T$ d'un filtre adapté à $s(t, a)$. Donc si l'on suppose que les

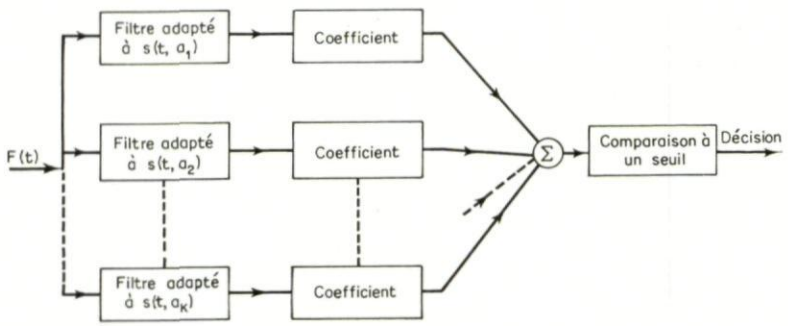


FIG. 8

valeurs possibles de a peuvent être représentées par une suite de valeurs discrètes, on peut toujours réaliser la décision précédente en construisant un certain nombre de filtres adaptés—un pour chaque valeur de a—multipliant les sorties par un coefficient convenable en faisant la somme et en comparant à un seuil (Fig. 8).

On peut également réaliser la règle de décision avec un ensemble de corrélateurs. Un exemple de réalisation de la figure 8 est l'ensemble de filtres étroits utilisés parfois dans les radars Döppler à impulsions. Mais le procédé devient rapidement impossible s'il y a plusieurs paramètres.

Donnons quelques exemples où la règle de décision prend une forme plus simple parce que l'intégration par rapport à a peut être effectuée explicitement.

3.1. *Signaux avec des Amplitudes Aléatoires*

L'exemple le plus simple de paramètre aléatoire est l'amplitude, c'est-à-dire:

$$s(t, a) = a\, s(t)$$

où a est une *constante pour chaque intervalle de détection* $O < t < T$ mais varie d'un intervalle à l'autre. Il en résulte une certaine densité de probabilité $p(a)$. Nous supposons, bien entendu, que toute l'information sur la présence ou l'absence de $s(t)$ est contenue dans l'intervalle $O < t < T$. Si d'une façon ou d'une autre, on sait quelque chose sur a, $p(a)$ pour l'intervalle considéré doit refléter cette connaissance. Nous considérons peutêtre une incertitude sur la surface effective, mais pas encore une cible "scintillante". Dans ce cas il vient:

$$\int_{-\infty}^{+\infty} p(a) \exp\left(\frac{a}{N_0}\int_0^T f(t)\, s(t)\, dt\right) \exp\left(-\frac{a^2}{2N_0}\int_0^T s^2(t)\, dt\right) da \gtrless \gamma \qquad (1)$$

Puisque le premier membre ne dépend de $f(t)$ que par:

$$L = \int_0^T f(t)\, s(t)\, dt$$

la quantité L est une statistique suffisante en ce sens que l'on résume en un seul nombre toute l'information utile.

Supposons maintenant que a ne puisse prendre que des valeurs positives. La dérivée du premier membre de Eq. (1) par rapport à L est:

$$\int_0^{\infty} \frac{a}{N_0} p(a)\, e^{(aL/N_0)-(a^2E/2N_0)}\, da > 0 \text{ qqs } p(a)$$

Le premier membre de Eq. (1) est monotone en L, donc:

$$L = \int_0^T f(t)s(t)\, dt \underset{NON}{\overset{OUI}{\gtrless}} \text{seuil}$$

Le "filtre adapté" reste une règle de décision optimum dans le cas d'une amplitude positive inconnue, qqs $p(a)$.

3.2. *Signal Fluctuant très Rapidement*

Pour une cible fluctuant très rapidement et un radar à impulsions, il se peut que l'amplitude et la phase de l'écho soient constantes pendant la durée de l'impulsion, mais fluctuent suffisamment rapidement pour admettre des valeurs indépendantes d'une impulsion à l'autre. En étendant l'exemple précédent, on arrive au système indiqué Fig. 9.

La structure de ce système est très proche de celle des radars à impulsions conventionnels, l'amplificateur IF correspondant au filtre et l'effet combiné du PPI et de l'observateur correspondant à l'opération échantillonnage et somme.

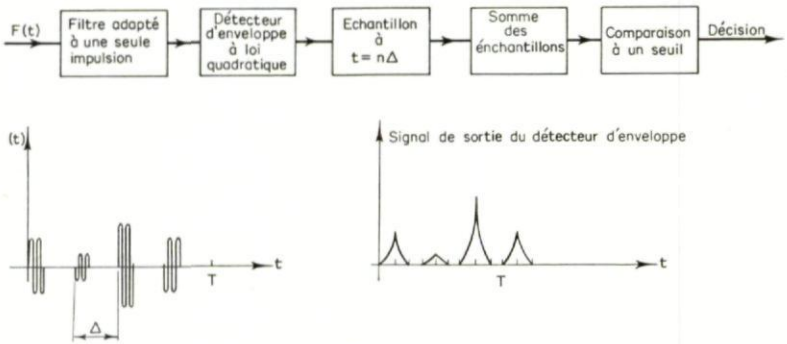


FIG. 9

Les performances de ce détecteur sont relativement simples à calculer. Les caractéristiques opératoires pour différents nombres d'impulsions sont indiquées sur les figures 10, 11, 12, and 13.

Supposons qu'un récepteur particulier reçoive une seule impulsion par balayage avec $Q_0 = 50$. Pour $Pf = 10^{-6}$ on trouve $P_D \sim 0{,}60$. Supposons que ce soit inacceptable et cherchons les différentes façons d'améliorer la situation sans toucher à la puissance crête. Multiplions la durée de l'impulsion par 3, on aura $Q_0 = 150$, $P_D \sim 0{,}83$. Une solution plus efficace peut être de diminuer la vitesse de balayage pour avoir trois impulsions avec des *amplitudes independantes* sur la même cible. $P_D = 0{,}96$ qui est infiniment meilleur. La raison en est, bien entendu, qu'il y a trois impulsions indépendantes, et que la chance qu'elles soient faibles toutes les trois est assez réduite. On peut montrer ce résultat d'une autre façon. Considérons un récepteur non optimum qui procède de la façon suivante: supposons qu'on reçoive trois impulsions dans l'intervalle $O < t < T$. Sur chaque impulsion on prend une décision sur la présence de la cible et finalement on annonce "cible présente" s'il y a eu détection sur l'une au moins des trois impulsions. Si P_D' et P_f' sont les probabilités pour une impulsion, il vient:

$$P_f = 1 - (1 - P_f')^3 = 3P_f' = 10^{-6}$$

$$\text{ou } P_f' = 3 \times 10^{-7}$$

Q_0 pour chaque inpulsion est 50, donc $P_D' = 0{,}56$ et:

$$P_D = 1 - (1 - 0{,}56)^3 = 0{,}915$$

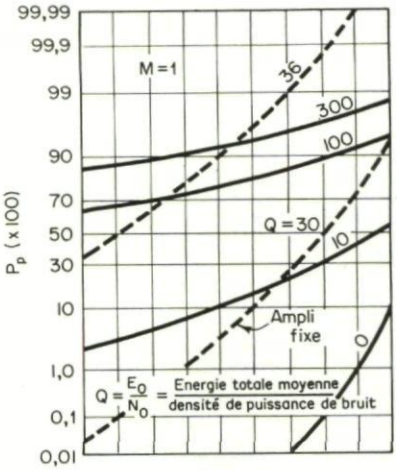


FIG. 10

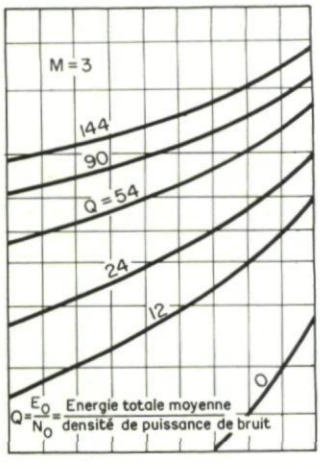


FIG. 11

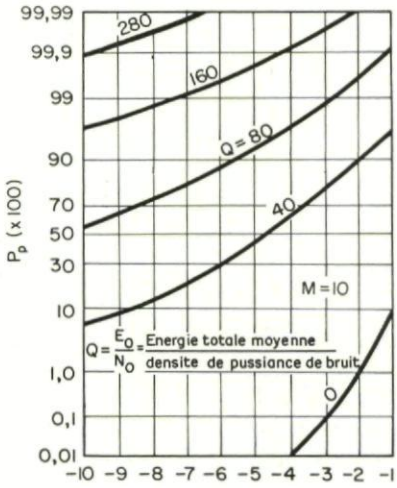
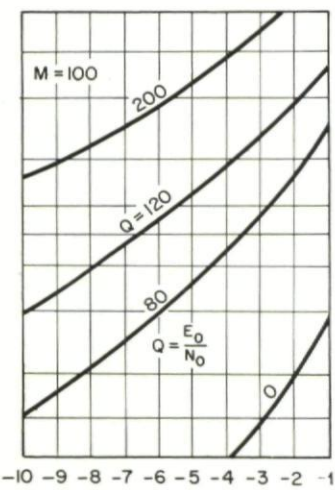


FIG. 12



$\log_{10} P_F$

FIG. 13

qui n'est pas tellement inférieur au système optimum qui donne 0,96. On utilise cette méthode d'intégration binaire en particulier pour le problème de l'intégration balayage par balayage avec des cibles fluctuant lentement, problème entièrement analogue à celui de l'intégration impulsion par impulsion avec des cibles fluctuant rapidement.

4. LE PROBLEME DE L'ESTIMATION

A part quelques cas particuliers de dispositifs d'alarme pure, la plupart des radars sont conçus pour mesurer certains paramètres tels la portée, la vitesse, l'azimuth . . . On peut même imaginer de mesurer les taux de

scintillation, l'accélération, la polarisation. Mais contentons-nous de la distance et de la vitesse. Woodward a posé les bases de toute une théorie sur la mesure simultanée de la vitesse et de la distance. C'est là qu'interviennent les qualités du signal autres que son énergie.

Supposons que le signal écho puisse se réprésenter par:

$$s(t, \tau', \omega', r, \Theta) = rE(t - \tau') \cos [(\omega_0 + \omega')t + \varphi(t - \tau') + \Theta]$$

t = temps
τ' = retard
ω' = pulsation Döppler
r = coefficient d'amplitude
Θ = coefficient de phase.

La méthode consiste à calculer la densité de probabilité a posteriori que le signal reçu F, contienne l'écho d'une cible caractérisée par des valeurs particulières des paramètres, d'intégrer cette densité par rapport à r' et Θ' et de trouver le maximum par rapport à τ' et ω'; ayant le rapport de vraisemblance $\Lambda(\tau', \omega')$. Supposons qu'un signal soit présent dans $f(t)$ avec le paramètre τ'', ω'' on montre que la forme de $\Lambda(\tau', \omega')$ quand une cible est présente à τ'',ω'' dépend de la fonction:

$$\Theta^2(\tau' - \tau'', \omega' - \omega'') = \left| \int_{-\infty}^{+\infty} u(t - \tau')u(t - \tau'') \, e^{j(\omega' - \omega'')^2} \right| dt$$

$u(t)$ = signal

fonction d'ambiguïté qui possède les propriétés suivantes:

(1) $\Theta^2(0, 0) \geq \Theta^2(\tau, \omega)$

(2) $\dfrac{1}{2\pi} \int_{-\infty}^{\infty} \int \Theta^2 (\tau, \omega) \, d\tau \, d\omega = 1$

La première propriété veut dire le maximum de $\Theta^2(\tau' - \tau'', -\omega' - \omega'')$ et par suite de $\Lambda(\tau',\omega')$ et de la probabilité a posteriori a lieu pour: $\tau' = \tau''$ $\omega' = \omega''$ la valeur juste. Mais la deuxième est plus intéressante et justifie le nom puisqu'elle implique une limitation à notre aptitude à faire une mesure simultanée de τ'' et ω''.

La largeur du maximum de $\Lambda(\tau', \omega')$ autour de $\tau' = \tau''$ $\omega' = \omega''$ détermine la précision de la mesure, et la présence d'autres maxima de valeurs comparables pour d'autres valeurs de τ' et ω', l'éventualité d'ambiguïté dans la mesure.

5. PROBLEMES DE LA DISCRIMINATION ET DU POUVOIR SEPARATEUR

5.1. *Généralités*

La principale application d'un radar de puissance est de surveiller un certain volume pour y détecter et poursuivre les cibles discrètes. Nous allons limiter notre discussion à cette application et aux radars monostatiques.

Un radar conçu pour surveiller un volume le fait en balayant (périodiquement ou non) l'angle solide correspondant. Les signaux diffusés par les cibles sont convertis après amplification éventuelle en hyperfréquence et subissent alors deux opérations:

(1°) par traitement des signaux reçus, l'information utile obtenue dans chaque balayage est extraite par filtrage, intégration, prise de décision.

(2°) par application de techniques totalement différentes, l'information extraite par la première opération est corrélée d'un balayage à l'autre pour former des trajectoires des cibles détectées.

Il paraît réaliste de considérer comme point de séparation entre les opérations l'endroit où a lieu une décision binaire sur la présence ou l'absence d'une cible dans chaque cellule de résolution avec estimation des paramètres qui caractérisent chaque cible (Fig. 14).

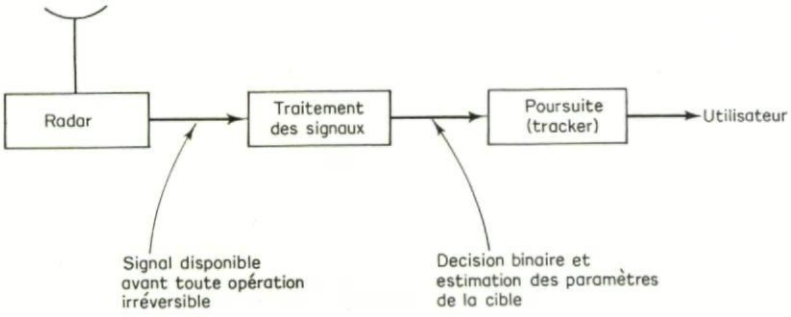


FIG. 14

Un certain nombre de considérations pratiques justifient cette division.

(1) Le niveau de performance demandé au radar pour produire une certaine qualité de poursuite peut être spécifié en termes de:
probabilité de détection
probabilité de fausse alarme
précision de l'estimation des paramètres.

(2) Les performances effectivement obtenues peuvent être appréciées objectivement à l'aide des mêmes quantités.

(3) C'est l'endroit où l'on peut facilement introduire un canal de communication à bande étroite si l'ensemble qui effectue la poursuite (tracker) n'est pas situé au même endroit que le radar.

(4) On peut combiner les données de plusieurs radars pour obtenir de meilleures performances.

(5) On peut introduire un opérateur.

5.2. *Conditions Imposées aux Qualités des Données par la Poursuite*

Le but est de démontrer un certain nombre de relations entre la qualité des données radar, mesurée par la probabilité de détection et fausse alarme et l'aptitude d'un radar à poursuivre une grande quantité de cibles.

Il est possible, en particulier, de mettre en évidence des dépendances entre l'aptitude au tracking et les probabilités de détection et de fausse alarme et de les relier au pouvoir séparateur du radar, sa période d'exploration, le nombre de cibles à poursuivre simultanément, à la logique du tracker, et à la mobilité des cibles.

Des résultats analogues s'appliquent en principe aux radars qui effectuent la poursuite pendant l'exploration.

On arrive aux résultats suivants:

(1) Il faut maintenir la probabilité de détection au-dessus d'une valeur minimum pour obtenir des performances raisonnables de poursuite.

(2) Il faut maintenir la probabilité de détection au-dessus d'une valeur minimum supérieure à la précédente pour démarrer rapidement les trajectoires sur de nouvelles cibles.

(3) La probabilité de fausse alarme doit rester faible pour ne pas perturber les trajectoires par de fausses cibles.

5.3. *Le Problème de la Discrimination*

Le radar reçoit, en plus des échos de cibles intéressantes, des échos d'autres cibles, du bruit interne, et du bruit de sources externes qui peut ne pas être du bruit blanc gaussien.

Pour discriminer entre les échos provenant des cibles utiles et tous les autres échos, il faut utiliser les caractéristiques des échos qui diffèrent substantiellement pour les deux catégories de cibles.

Ces caractéristiques sont:

(1) l'instant d'arrivée (distance)

(2) l'angle d'arrivée (azimuth, élévation)

(3) coefficients de la matrice de diffusion

(4) structure fine

(5) variation avec le temps des quantités précédentes.

(a) Les caractéristiques à utiliser pour une application donnée résultent de l'examen de la situation. La distance et l'angle d'arrivée sont presque toujours utilisés pour fournir une discrimination partielle.

(b) La matrice de diffusion offre des possibilités qui n'ont pas été complètement exploitées. La matrice de diffusion relie les composantes du vecteur réfléchi par la cible à celles du vecteur émis.

$$\begin{bmatrix} E_x^{\,r} \\ E_y^{\,r} \end{bmatrix} = \begin{bmatrix} a_{11} \exp(i\varphi_{11}) & a_{12} \exp(i\varphi_{12}) \\ a_{21} \exp(i\varphi_{21}) & a_{22} \exp(i\varphi_{22}) \end{bmatrix} \begin{bmatrix} E_x^{\,t} \\ E_y^{\,t} \end{bmatrix}$$

$$a_{12} = a_{21} \quad \varphi_{12} = \varphi_{21}$$

Exemple :

Discrimination contre les gouttes de pluie : par renversement du sens de polarisation d'une onde polarisée circulairement.

De nombreuses mesures ont été faites sur certains facteurs de la matrice, et montrent que dans une certaine mesure on peut classer les cibles par leurs propriétés de polarisation, mais cette technique est difficile à utiliser.

On utilise toutefois l'amplitude d'un des termes de la matrice, plus la distance et l'angle d'arrivée pour discriminer entre les avions et les oiseaux sur les aérodromes.

(c) La structure fine est la modulation a l'intérieur de l'impulsion des éléments de la matrice et semble offrir des possibilités exceptionnelles mais difficiles a exploiter de discrimination.

(d) Variation avec le temps des quantités précédentes.

La plus utilisée est la phase à cause des vitesses des cibles.

Si l'on considère seulement l'amplitude et la phase d'un élément de la matrice de diffusion, la tension du signal écho provenant d'une cible éclairée par un radar a impulsions s'écrit :

$$u(t) = \sum_{n=0}^{n=N-1} h\left(t - hT_1 - \frac{2R}{c}\right) \exp\left[i\,\omega_0\left(t - \frac{2R}{c}\right)\right]$$

($u(t)$ dépend également de la distance et de la période d'exploration).

$h(t) = 1$ pour $0 < t < \tau$ durée de l'impulsion

$h(t) = 0$ ailleurs

T_1 période entre impulsions

$(N - 1)\,T_1$ est le temps pendant lequel la cible est éclairée.

Si R s'écrit :

$$R = R_0 + \frac{dR}{dt}\,t$$

où R_0 est la distance à $t = 0$, la phase φ de l'écho devient :

$$\varphi = \frac{2\omega_0}{c}\left(R_0 + \frac{dR}{dt}\,t\right)$$

$$\frac{2\omega_0}{c}\,R_0 = \text{déphasage constant}$$

$$\frac{2\omega_0}{c}\,\frac{dR}{dt}\,t = \omega_D t$$

phase variable linéairement avec la fréquence Döppler $\omega_D/2\pi$ proportionnelle à la vitesse radiale :

$$V_r = \frac{dR}{dt}$$

Tout procédé de discrimination fondé sur la fréquence Döppler est sujet à des contraintes fondamentales provenant de la fonction d'ambiguïté de

Woodward, du spectre Döppler des signaux radars eux-mêmes et de la perturbation des échos par le bruit.

Illustrons ces problèmes.

La fonction d'ambiguïté de Woodward s'écrit pour un signal $u(t)$:

$$X(\tau, \omega) = \int_{-\infty}^{+\infty} u(t)\, u^*(t + \tau) \exp(-i\omega t)\, dt$$

où τ est le retard et ω la pulsation Döppler.

Une interprétation de cette fonction est que si deux cibles sont présentes simultanément aux points τ_1, ω_1 et τ_2, ω_2 la quantité :

$$|X(\tau_2 - \tau_1, \omega_2 - \omega_1)|$$

est une mesure du pouvoir séparateur du radar. Une valeur faible correspond à un bon pouvoir séparateur, tandis que si cette quantité est voisine de 1 pour n'importe quel couple de points du plan (τ, ω) le pouvoir séparateur est nul.

Le problème le plus intéressant est celui de la séparation d'une cible parmi un grand nombre d'autres cibles : on intègre $|X(\tau, \omega)|^2$ sur l'ensemble des valeurs possibles de (τ, ω), chaque valeur étant affectée d'un coefficient fonction de (τ, ω) qui représente la probabilité d'apparition d'une cible au point (τ, ω). Le rapport de cette intégrale pour une cible intéressante à l'intégrale correspondante pour n cibles sans intérêt, pourrait être utilisé comme une mesure du pouvoir séparateur.

Malheureusement, on connaît mal les coefficients à appliquer. Les coefficients correspondent aux distributions probables en distance, vitesse et accélération des cibles intéressantes et des autres.

Considérons le cas où $u(t)$ est une impulsion unique, non modulée, de durée T et d'amplitude unité. La fonction d'ambiguïté pour le retard τ, la pulsation Döppler ω et l'accélération : $\alpha = d\omega/dt$ (proportionnelle à accélération radiale) s'écrit :

$$X(\tau, \omega, \alpha) = \frac{\pi}{\alpha} \left| \int_{\sqrt{\alpha/\pi}(\omega/\alpha)}^{\sqrt{\alpha/\pi}[(\omega/\alpha)+T-\tau]} \exp\left(-\frac{i\pi}{2}\right) V^2\, dV \right|$$

On obtient des surfaces d'ambiguïté que l'on peut représenter par les intersections avec les plans $\alpha = 0$, $\omega = 0$, $\tau = 0$.

On constate que dans le cas où il n'y a pas de bruit, deux cibles seront complètement séparées en distance si :

$$\frac{\tau}{T} > 1 \text{ ou } \frac{2R_0}{cT} > 1$$

ce qui était évident a priori. Si l'on veut séparer des cibles plus proches, il faut utiliser une modulation à l'intérieur de l'impulsion.

Le pouvoir séparateur pour la vitesse est plus difficile à mettre en évidence. Toutefois la largeur du maximum principal est :

$$\frac{\omega T}{2\pi} = 2 \text{ ou } \omega = \frac{4\pi}{T}$$

Si on utilise cette quantité comme mesure du pouvoir séparateur, on obtient:

$$\Delta V_r = \frac{\lambda_0}{T}$$

Bien entendu, aucun radar n'émet d'impulsion unique non modulée. Le pouvoir séparateur indiqué par les considérations ci-dessus est celui que fournit une impulsion prise dans le train d'impulsions.

Nous allons tout d'abord reconsidérer ces problèmes de précision sur les mesures des paramètres et de pouvoir séparateur en présence de bruit.

5.4. *Détection de Signaux avec des Paramètres Inconnus*

La modulation du signal est de la forme:

$$u(t; a_1, a_2 \cdots a_N) = E(t; a_1, a_2 \cdots a_N) \exp \{j\Phi(t; a_1 \cdots a_N)\}$$

avec:

$$\int_{-\infty}^{+\infty} |u(t; a_1 \cdots a_N)|^2 \, dt = 1$$

de telle sorte qu'un signal dont la modulation est:

$$Ae^{j\delta} u(t; a_1 \cdots a_N)$$

contient une énergie $E_s = A^2/2$.

Il faut calculer les corrélations:

$$\xi (a_1 \cdots a_N) = \int_{-\infty}^{+\infty} u(t; a_1 \cdots a_N)^* \, W(t) \, dt$$

et chercher le maximum de $|\xi(a_1 \cdots a_N)|$ pour l'ensemble des paramètres.

Supposons que la modulation d'entrée $W(t)$ se décompose en un terme signal:

$$Ae^{j\delta} u(t; \beta_1 \cdots \beta_N)$$

les β_i étant les valeurs des paramètres et un terme bruit $n(t)$.

Alors $\xi(a_1 \cdots a_N)$ se décompose:

$$\xi(a_1 \cdots a_N) = Ae^{j\delta} \int_{-\infty}^{+\infty} u(t; a_1 \cdots a_N)^* \, u(t; \beta_1, \cdots \beta_N) \, dt + \int_{-\infty}^{+\infty} u(t; a_1 \cdots a_N)^* \, n(t) \, dt$$

La fonction de corrélation du signal, soit Γ, est définie par:

$$\Gamma(a_1 \cdots a_N; \beta_1 \cdots \beta_N) = \int_{-\infty}^{+\infty} u(t; a_1 \cdots a_N)^* \, u(t; \beta_1 \cdots \beta_N) \, dt$$

et le critère de détection peut s'écrire:

$$\operatorname*{Max}_{a_1 \cdots a_N} | Ae^{j\delta} \, \Gamma(a_1 \cdots a_N; \beta_1 \cdots B_N) + \zeta(a_1 \cdots a_N)| \geq K$$

où:

$$\zeta = \int_{-\infty}^{+\infty} u(t; a_1 \cdots a_N)^* \, n(t) \, dt$$

On peut montrer que la fonction de corrélation complexe pour ces termes de bruit pour différents ensembles de paramètres est donnée par:

$$E[\zeta(a_1 \cdots a_N)\ \zeta^*(\beta_1 \cdots \beta_N)] = 4N_0\Gamma(a_1 \cdots a_N; \beta_1 \cdots \beta_N)$$

$E[\ \ ]$ = Espérance mathématique.

N_0 densité spectrale de bruit au voisinage de f_0.

Le problème de détection et de précision est donc complètement déterminé mathématiquement par la fonction de corrélation du signal.

Remarquons que:

$$\Gamma(a_1 \cdots a_N; a_1 \cdots a_N) = 1$$

tandis que l'inégalité de Schwarz montre que:

$$|\Gamma(a_1 \cdots a_N; \beta_1 \cdots \beta_N)| \leq 1$$

Si $\Gamma(a_1 \cdots a_N; \beta_1 \cdots \beta_N)$ est voisin de l'unité pour des ensembles $(a_1 \cdots a_N)$ différents de $(\beta_1 \cdots \beta_N)$, il y a une ambiguïté dans l'estimation des paramètres inhérente au signal émis particulier.

Supposons que le rapport signal sur bruit E_s/N_0 soit grand et que la fonction de corrélation du signal n'ait pas de maxima secondaires appréciables, de telle sorte que les estimations $(\hat{a}_1 \cdots \hat{a}_N)$ des paramètres soient proches des valeurs vraies $(\beta_1 \cdots \beta_N)$. Pour étudier la précision fournie par le signal émis combiné à notre dispositif de réception (filtre adapté généralisé), développons les fonctions:

$$\Gamma(a_1 \cdots a_N; \beta_1 \cdots \beta_N) \text{ et } \zeta(a_1 \cdots a_N)$$

en séries de Taylor multiples autour du point:

$$a_n = \beta_n \quad n = 1 \cdots N$$

En conservant seulement les termes du deuxième ordre nous pouvons résoudre explicitement par rapport aux estimateurs $(\hat{a}_1 \cdots \hat{a}_N)$ obtenant des approximations valables au premier ordre près dans le bruit. De cette façon, les quantités $(\hat{a}_1 \cdots \hat{a}_N)$ sont des variables aléatoires gaussiennes dont les moyennes sont les valeurs vraies et qui sont caractérisées par une matrice de termes:

$$M_{Ke} = E[(\hat{a}_K - \beta_K)(\hat{a}_e - \beta_e)]$$

qui donnent les variantes des erreurs et les corrélations entre elles. Cette matrice est déterminée par le signal. En fait l'inverse M^{-1} est donnée par:

$$(M^{-1})_{Ke} = -\frac{E_S}{2N_0}\left[\frac{\partial^2}{\partial a_K \partial a_e} Q(a_1 \cdots a_N; \beta_1 \cdots \beta_N)\right]_{\cdots\cdots}$$

$$(a_1 \cdots a_N) = (\beta_1 \cdots \beta_N)$$

En termes de la fonction d'ambiguïté:

$$Q(a_1 \cdots a_N; \beta_1 \cdots \beta_N) = |\Gamma(a_1 \cdots a_N; \beta_1 \cdots \beta_N)|^2$$

Ce processus d'approximation a une interprétation géométrique simple: le maximum principal de la fonction d'ambiguïté (c'est-à-dire la région près du point $a_n = \beta_n$) est approchée par une forme quadratique des différences $a_n - \beta_n$. Cette forme quadratique est un paraboloïde à N dimensions dont le maximum est égal à 1 au point:

$$a_n = \beta_n \quad n = 1 \cdots N$$

et qui s'adapte à la courbure de la surface d'ambiguïté en ce point.

5.5. *Précision dans la Mesure de la Distance, de l'Azimuth et de la Vitesse*

Nous allons considérer des signaux qui ont deux paramètres inconnus: le retard τ, la fréquence Döppler f, en plus de l'amplitude et de la phase de la porteuse.

La fonction de modulation du signal dépend donc de deux paramètres:

$$u(t; \tau, f) = u(t - \tau)\, e^{2\pi j f(t-\tau)}$$

$$\text{où} \quad u(t) = E(t)\, e^{j\Phi(t)}$$

est proportionnel au signal émis.

Nous supposons que:

$$\int_{-\infty}^{+\infty} |u(t)|^2\, dt = 1$$

L'effet du retard τ sur la porteuse a été inclus dans la phase inconnue de la porteuse et l'effet Döppler sur la fonction $u(t)$ elle-même a été négligé (en général très faible par rapport à l'effet Döppler sur la porteuse).

Les corrélations à effectuer sur la modulation d'entrée $W(t)$ sont de la forme:

$$\xi(\tau, f) = \int_{-\infty}^{+\infty} u^*(t - \tau)\, e^{-2\pi j f(t-\tau)}\, W(t)\, dt$$

Ce qui revient à déplacer la fréquence de la porteuse modulée par $W(t)$ d'une quantité f et à faire passer le résultat dans un filtre adapté. La sortie passe dans un détecteur d'enveloppe, ce qui fournit: $|\xi(\tau, f)|$. Ceci peut être effectué pour un ensemble de fréquences f, tandis que τ apparaît comme un retard dans le signal de sortie du filtre. Les détails de cette opération sont donnés plus loin. Nous nous intéressons à la précision maximum que l'on peut atteindre avec un filtre adapté parfait.

Pour la famille de signaux $u(t; \tau, f)$ la fonction de corrélation est de la forme:

$$\begin{aligned} Q(t, f, \tau', f') &= \int_{-\infty}^{+\infty} u^*(t - \tau) u(t - \tau')\, e^{2\pi j[f'(t-\tau') - f(t-\tau)]} \\ &= e^{2\pi j f'(\tau - \tau')} \int_{-\infty}^{+\infty} u^*(t) u(t + \tau - \tau')\, e^{-2\pi j(f - f')t}\, dt \end{aligned}$$

Nous allons utiliser la fonction:

$$\int_{-\infty}^{+\infty} u^*(t)\, u(t + \tau - \tau')\, e^{-2\pi j(f-f')t}\, dt$$

comme fonction de corrélation du signal. Il est très important que cette fonction ne dépende que des différences entre les paramètres. Pour mettre ce point en évidence, écrivons:

$$\chi(\tau,f) = \int_{-\infty}^{+\infty} u^*(t)u(t + \tau)\, e^{-2\pi jft}\, dt$$

La fonction d'ambiguïté est:

$$Q(\tau, f\,;\tau', f') = |\Gamma(\tau, f\,;\tau^1, f')|^2$$

qui ne dépend que des différences. Posons:

$$Q(\tau, f\,;\tau', f') = \Psi(\tau - \tau', f - f')$$

avec:

$$\Psi(\tau, f) = \left| \int_{-\infty}^{+\infty} u^*(t)u(t + \tau)\, e^{-2\pi jft}\, dt \right|^2$$

D'après le paragraphe précédent, pour évaluer les erreurs dans l'estimation de τ et de f, il faut dériver $Q(\tau, f\,;\tau', f')$ par rapport à τ et f, évaluer ces dérivés pour $\tau = \tau', f = f'$ où τ', f' représentent les valeurs vraies. Il vient par exemple:

$$\left[\frac{\partial^2}{\partial\tau^2} Q(\tau, f\,;\tau', f')\right]_{\substack{\tau=\tau' \\ f=f'}} = \frac{\partial^2}{\partial\tau^2} \Psi(\tau, f)_{\tau=0}$$

etc . . . qui ne dépend pas des valeurs vraies τ', f', ce qui veut dire que le erreurs sur les mesures par radar du retard et de la fréquence Döppler, ainsi que la corrélation entre ces erreurs sont indépendantes des valeurs des quantités mesurées. Ce résultat qui implique que l'erreur sur la distance est indépendante de la distance (à rapport signal sur bruit constant) provient entièrement du fait que la fonction d'ambiguïté ne dépend que des différences des paramètres.

Introduisons la transformée de Fourier $U(f)$ de la modulation $u(t)$:

$$U(f) = \int_{-\infty}^{+\infty} u(t)\, e^{-2\pi jft}\, dt$$

En termes de $U(f)$ il vient:

$$\Psi(\tau, f) = \left| \int_{-\infty}^{+\infty} U^*((\nu)\; U(\nu + f)\, e^{2\pi j\nu t}\, d\nu \right|^2$$

Si nous calculons les dérivées pour obtenir la matrice M^{-1}, puis M par inversion, nous aurons les formules d'erreurs cherchées.

Posons:

$$T^2 = \int_{-\infty}^{+\infty} |u(t)|^2 t^2 \, dt - \left[\int_{-\infty}^{+\infty} |u(t)|^2 t \, dt \right]^2$$

$$F^2 = \int_{-\infty}^{+\infty} |U(\nu)|^2 \nu^2 \, d\nu - \left[\int_{-\infty}^{+\infty} |U(\nu)|^2 \nu \, d\nu \right]^2$$

$$\rho = \frac{1}{2\pi FT} \left\{ \int_{-\infty}^{+\infty} |u(t)|^2 t \, \Phi'(t) \, dt - \int_{-\infty}^{+\infty} |u(t)|^2 t \, dt \right.$$
$$\left. \times \int_{-\infty}^{+\infty} |u(t)|^2 \Phi'(t) \, dt \right\}$$

où $\Phi'(t)$ est la dérivée de la modulation de phase $\Phi(t)$. Par rapport à ces quantités il vient:

$$\overline{(\Delta\tau)^2} = \frac{1}{(E_s/N_0)4\pi^2F^2(1 - \rho^2)} \tag{2}$$

$$\overline{(\Delta f)^2} = \frac{1}{(E_s/N_0)4\pi^2T^2(1 - \rho^2)} \tag{3}$$

et l'on a:

$$\overline{\Delta\tau \, \Delta f} = \rho \left[\overline{(\Delta\tau)^2} \, \overline{(\Delta f)^2} \right]^{1/2}$$

Si l'on considère $|u(t)|^2$ comme une densité de probabilité, T^2 a le caractère d'une variance. Donc T est une mesure de la durée du signal; de même F est une mesure de sa largeur de bande.

Les formules (2) et (3) donnent les erreurs sur la mesure de τ et de f quand ces quantités sont mesurées simultanément. L'erreur dans la mesure du retard quand f est connu, s'obtient en faisant $\rho = 0$ dans (2). Si nous appelons ces erreurs:

$$\overline{(\Delta\tau)_0^2} \text{ et } \overline{(\Delta f)_0^2}$$

il vient:

$$\overline{(\Delta\tau)_0^2} = \frac{1}{(E_s/N_0)4\pi^2F^2} \tag{4}$$

$$\overline{(\Delta f)_0^2} = \frac{1}{(E_s/N_0)4\pi^2T^2} \tag{5}$$

Les formules (2) à (5) montrent quantitativement le résultat bien connu qu'une mesure précise de la fréquence nécessite un temps d'observation assez long, et une mesure précise du retard, une largeur de bande assez grande.

Introduisons les quantités:

$$T_{\text{eff}} = \text{durée effective du signal} = T(1 - \rho^2)^{1/2} \tag{6}$$

$$F_{\text{eff}} = \text{largeur de bande effective du signal} = F(1 - \rho^2)^{1/2} \tag{7}$$

La quantité qui détermine la précision à deux dimensions d'un signal est son produit durée—largeur de bande $T_{\text{eff}} F_{\text{eff}}$ qui est différent de TF.

En termes des quantités physiques, vitesse V et distance R, les formules deviennent:

$$\boxed{\delta_V = \frac{\lambda_0}{4\pi\sqrt{E_s/N_0}\, T\sqrt{1-\rho^2}}} \tag{8}$$

Exemple: pour une impulsion de durée δ, en l'absence de corrélation, on trouve:

$$T = \frac{\delta}{2\sqrt{3}}$$

soit:

$$\delta_V = \frac{\sqrt{3}\,\lambda_0}{2\pi\sqrt{E_s/N_0}\,\delta}$$

formule valable également pour un train d'impulsions très courtes de durée globale δ.

De meme

$$\boxed{\delta R = \frac{c}{4\pi\sqrt{E_s/N_0}\, F\sqrt{1-\rho^2}}} \tag{9}$$

La mesure de la distance est "incohérente" en ce sens que la distance est déterminée par l'effet du retard de propagation sur la modulation, sans s'occuper de la phase de la porteuse. Tandis que la mesure de la vitesse est cohérente et est déterminée par l'effet du déplacement de la cible sur la phase de la porteuse. Si l'on effectue deux mesures incohérentes R_1, R_2, séparées de T_1 secondes, on obtient une mesure incohérente de la vitesse par:

$$V = \frac{R_2 - R_1}{T_1}$$

la précision est:

$$\delta_V' = \sqrt{2}\,\frac{\delta R}{T_1}$$

ou si $\rho = 0$

$$\delta_V' = \boxed{\frac{\sqrt{2}\; c/F}{4\pi\sqrt{E_s/N_0}\; T_1}}$$

On peut traiter à l'aide de ces théories les problèmes de pouvoir séparateur pour deux cibles, puis pour une multitude de cibles, et le problème de la synthèse du meilleur signal pour la séparation d'une multitude de cibles.

5.6. *Discrimination Entre Échos Radars et Bruit*

Le bruit peut être interne ou externe.

Le bruit interne provient de l'agitation thermique et du "shot effect." Il est caractérisé avec un excellent degré d'approximation par un processus

gaussien aléatoire stationnaire, ce qui permet en ignorant le bruit externe de résoudre les problèmes de détection et d'estimation des paramètres.

Le bruit externe provient de sources très diverses telles que, sources cosmiques, autres radars, émetteurs de trafic, radio-altimètres et pour les radars militaires une grande variété de brouilleurs.

Les distributions spectrales peuvent être quelconques, depuis la distribution uniforme jusqu'à la fonction δ. Une variété infinie de distributions d'amplitude est également possible.

Pour essayer de traiter le problème du bruit externe, nous allons considérer séparément les effets de diverses distributions spectrales, puis de diverses distributions d'amplitude, ce qui est évidemment un artifice puisqu'elles ne sont pas indépendantes.

Ceci va nous permettre de décrire et de justifier deux méthodes de protection contre les bruits externes.

Considérons un radar cohérent qui transmet le train d'impulsions:

$$S_t(t) = \sum_{N=0}^{N-1} h(t - nT_1) \exp i\omega_0(t - nT_1)$$

où $h(t) = 1$ pour $0 < t < \tau$

$h(t) = 0$ ailleurs

T_1 = période entre impulsions

τ = longueur de l'impulsion

ω_0 = fréquence porteuse

N = le nombre d'impulsions qui éclairent la cible.

Pour une cible à la distance R_0 qui se déplace avec une vitesse radiale uniforme dR/dt le signal reçu en négligeant le bruit pour le moment:

$$Sr(t) = \sum_{n=0}^{n=N-1} r_n h\left(t - nT_1 - \frac{2R_0}{c}\right) \times \exp\left[i\omega_0\left(t - nT_1 - \frac{2R_0}{c} - \frac{2}{c}\frac{dR}{dt}nT_1\right)\right]$$

où r_n représente l'amplitude de la $n^{\text{ième}}$ impulsion. On a fait plusieurs approximations:

(1) $\frac{2}{c}\frac{dR}{dt}(N-1)T_1 \ll \tau$ donc ce terme a été négligé dans l'argument de $h(t)$.

(2) En écrivant $\omega_D = \frac{2\omega_0}{c}\frac{dR}{dt}$ on suppose que $\omega_0\tau$ est très faible, c'est-à-dire que la fréquence Döppler est suffisamment faible et l'impulsion suffisamment courte pour qu'il y ait une très petite fraction de cycle Döppler à l'intérieur de l'impulsion.

(3) Les facteurs phases qui ne changent pas d'une impulsion à l'autre ont été négligés.

Hypothèses et approximations valables pour une large catégorie de radar à impulsions.

Supposons que l'on fasse passer le signal reçu à travers une ligne à retard à plots telle que l'indique la figure 15. La $n^{\text{ième}}$ impulsion est donc retardée de $(N-1-n)\,T_1$, fixons notre attention sur l'intervalle:

$$(N-1)\,T_1 + \frac{2R_0}{c} < t < (N-1)\,T_1 + \frac{2R_0}{c} + \tau$$

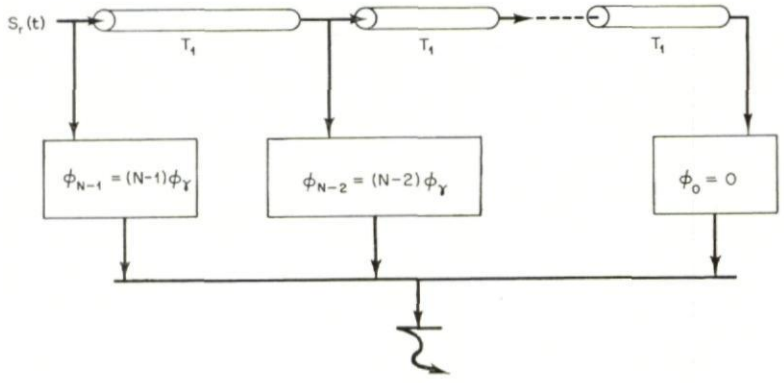


FIG. 15

Nous trouvons que le signal qui apparaît sur le $n^{\text{ième}}$ plot de la ligne (numérotés depuis la droite) s'écrit:

$$S_{rn}(t) = r_n \exp\left(-i\omega_D\, nT_1\right) \exp\left\{i\omega_0 \left[t - (N-1)\;T_1 - \frac{2R_0}{c}\right]\right\}$$

puisque le facteur qui représente la porteuse est le même à tous les plots, considérons seulement les facteurs:

$$r_n \exp\left(-i\omega_D nT_1\right)$$

Si le récepteur a une largeur de bande égale à $1/\tau$ le bruit reçu peut être représenté à chaque plot de la ligne à retard par:

$$a_n \exp i\,\Psi_n$$

Un procédé de détection consiste à introduire un déphasage φj à chaque plot, d'additionner les échantillons résultant (signal et bruit) linéairement et de comparer la somme à un seuil C. Si:

$$\left|\sum_{n=0}^{N-1} \{r_n \exp\left[in(\varphi_j - \omega_D T)\right] + a_n \exp\left[i(\Psi_n + n\varphi_j)\right]\}\right| \geqslant C \qquad (10)$$

on dira qu'il y a cible. Si $a_n \exp i\Psi_n$ représentent des échantillons d'un processus aléatoire connu, stationnaire et gaussien, la sensibilité maximum aura lieu pour:

$$\varphi_j = \omega_D T$$

Comme en général, on ne connaît pas ω_D à l'avance, il est nécessaire d'avoir une série de déphaseurs à chaque plot pour couvrir toutes les fréquences Doppler possibles.

On montre que le critère de détection défini par la relation (10) est optimum, en ce sens que pour une probabilité de fausse alarme donnée il fournit la plus forte probabilité de détection, dans l'hypothèse où les $a_n \exp i\Psi_n$ représentent un processus gaussien. Il est intéressant de supprimer cette restriction.

Rappelons qu'il est important d'obtenir la sensibilité maximum, mais qu'il est primordial de contrôler le taux de fausse alarme. Considérons tout d'abord un procédé qui permet d'obtenir ce résultat.

Si on suppose $r_n = 0$ quel que soit n, c'est-à-dire s'il n'y a pas de signal utile, la quantité sur laquelle est fondée la détection (10) peut devenir très grande avec des Ψ_n en progression uniforme, ce qui peut arriver par exemple si le bruit contient un signal CW a l'intérieur de la bande passante du récepteur. Le résultat serait un taux de fausse alarme excessif (brouillage). Un résultat analogue si la distribution d'amplitude des a_n avait une densité assez forte pour les fortes valeurs de a.

Supposons que le signal émis contienne un facteur de phase exp $(-i\Theta_n)$ où Θ_n est une variable aléatoire uniformément distribuée sur l'intervalle $0.2\ \pi$ le signal émis devient:

$$S_t(t) = \sum_{n=0}^{N-1} h(t - nT_1) \exp \{i[\omega_0(t - nT_1) - \Theta_n\}]$$

Le facteur de phase est mis en mémoire pendant une période T_1, et réintroduit dans le signal reçu avec un signe opposé, les autres opérations restant les mêmes, le critère de détection devient:

$$\left| \sum_{n=0}^{N-1} \{r_n \exp [in(\varphi_j - \omega_D T)] + a_n \exp [i(\Psi_n + n\varphi_j + \Theta_n)]\} \right| > C$$

Le résultat est de laisser intact le signal et de rendre aléatoire la phase du bruit, quel que soit le degré de régularité ou d'irrégularité qu'avait auparavant la phase du bruit l'addition des Θ_n assure que:

$$\exp [i(\Psi_n + n\Psi_j + \Theta_n)]$$

est une variable aléatoire. Donc, quand les composantes de bruit externe sont additionnées dans les réseaux déphaseurs, on ne peut plus les distinguer du bruit aléatoire par leur distribution spectrale. Le probabilité de fausse alarme est donc contrôlée.

On peut utiliser la fréquence au lieu de la phase.

Une fois qu'on a rendu la fréquence ou la phase aléatoire pour contrôler le taux de fausse alarme, il reste le problème de l'infinité de distribution des a_n.

Il n'est évidemment pas possible d'envisager toutes les situations mais on peut utiliser des modèles simples de simulation.

CHAPTER 6

TARGET RESOLUTION: CAPABILITIES OF MODERN RADAR AND FUNDAMENTAL LIMITS

A. W. Rihaczek*

Aerospace Corporation, El Segundo, California, U.S.A.

The growth in radar requirements, from crude short-range measurements to dense-target resolution at long ranges, is paralleled by a corresponding increase in the sophistication of radar systems: from simple constant-carrier pulses to post-detection integration, coherent integration of pulse trains, pulse compression signals, and the coherent processing of trains of such signals. The present study starts with an interpretation of the radar uncertainty relation in its significance for target resolution, showing the role of waveform designs as a means of achieving a match between the transmitted signal and the characteristics of the target environment. This provides a framework into which the various principles of high-resolution radar are fitted. In discussing the limitations on resolution performance, it is shown that achievable target resolution depends on the characteristics of the target environment in which radar operates, the number of targets, and the size of the delay-Doppler space they occupy. These findings are applied to two practical examples: (1) *the ground mapping radar using the synthetic aperture principle and* (2) *the case of extended target "clouds" consisting of a large number of discrete scatterers.*

1. INTRODUCTION

From the basic task of radar, the detection of a target and determination of its range, the step toward performing the same measurements on a group of targets is seemingly a small one; however, this step has so far taken the effort of the past two decades and is not yet completed. In the development of radar, multiple target resolution has turned out to be one of the most difficult problems to solve. The following discussion is an attempt to expose the fundamental nature of the resolution problem and, in putting the modern radar techniques into proper context, to present a unified view of the capabilities and limitations of radar for target resolution.

The trivial question of resolution when targets have an angular separation of more than a beamwidth will not be considered here. Moreover, it will be assumed that all targets actually illuminated by the beam are being illuminated with the same signal strength, so that no resolution on the basis

* The study was supported by the Air Force Space System Division under Contract No. AF 04(695)-269.

of the beam pattern is possible. Thus we shall study the problem of target resolution on the basis of range and range rate measurements. Range acceleration and higher order range derivatives will be neglected, which fact is not too restrictive in radar practice as the effect of target acceleration, and more so of higher order range derivatives, usually is indeed negligible over the coherent signal processing time.

It will be further assumed that the radar uses a matched-filter or correlation receiver; that is, a receiver which gives optimum detection performance in additive, white Gaussian noise. An obvious justification for this assumption is the fact that most target detection radars use this type of receiver or at least approximate it. However, even though target resolution appears to be the problem of signal detection in a signal background rather than a background of thermal noise, it is not likely that a more efficient receiver than the matched-filter type can be found. A justification of this contention is beyond the scope of this paper, but can be indicated by pointing out that the essence of target resolution is in concentrating the individual returns in delay and Doppler so that they can be recognized as separate returns. Allowing a suitable modification of the transmitted signal waveform, this goal can be best achieved with a matched-filter system, and the problem becomes that of signal design for optimum resolution performance.

The basis for analyzing combined range and range rate resolution in a matched-filter radar was laid by Woodward[1] through his study of the combined filter response in delay and Doppler or so-called ambiguity function. Work along the same lines was continued by Siebert[2] and an application of these concepts to target resolution in clutter was given by Westerfield, et al.[3], and to resolution in dense-target environment by Fowle, et al.[4] The purpose of the present study is to give a general exposition of the resolution problem, so as to show what high-performance radar can and cannot do.

2. AMBIGUITY FUNCTION AND RADAR UNCERTAINTY RELATION

First, we summarize briefly the notation and basic concepts to be used in the following analysis of target resolution. Signals will be written in the complex notation as introduced by Gabor[5]

$$\psi(t) = \mu(t)\, e^{j2\pi f_0 t} = a(t)\, e^{j\phi(t)}\, e^{j2\pi f_0 t}\ , \tag{1}$$

where $\mu(t)$ is the complex modulation function of the real signal, f_0 is the carrier frequency, $a(t)$ is the amplitude modulation function or real signal envelope, and $\phi(t)$ is the phase modulation function. The frequency modulation function is the time-derivative of the phase modulation function, $\phi'(t)$. The real signal is the real part of the complex signal, and the complex signal is obtained from the real signal by omitting the negative frequencies and doubling the amplitude of the positive frequencies. In the case of radar, the signal bandwidth is usually small compared to the carrier frequency, and the modulation functions defined above have practical meaning.

The notation will be simplified without any restrictions on the generality of the results by normalizing the signal energy such that

$$\int_{-\infty}^{+\infty} |\mu(t)|^2 \, dt = \int_{-\infty}^{+\infty} |M(f)|^2 \, df = 1 \ , \tag{2}$$

where $M(f)$ is the Fourier transform of the complex modulation function $\mu(t)$, that is, the frequency spectrum of the complex modulation function. A further simplification is achieved by defining the carrier frequency, in agreement with common usage of the term, as the normalized (because of Eq. 2) first moment of the energy density spectrum of the signal,

$$f_0 = \int_0^{\infty} f \, |\Psi(f)|^2 \, df \ , \tag{3}$$

and choosing the origin of time such that the normalized first moment of the squared signal envelope is zero,

$$\bar{t} = \int_{-\infty}^{+\infty} t \, |\mu(t)|^2 \, dt = 0 \ . \tag{4}$$

With the above conventions, the definitions for three very useful signal parameters take relatively simple forms. The first parameter, the effective signal bandwidth β_0, plays an important role for measurement precision and resolution in range. It is defined as the rms deviation of the energy density spectrum from its mean or β_0^2, equivalently, as the second moment of the energy density spectrum of the modulation function:

$$\beta_0^2 = (2\pi)^2 \int_{-\infty}^{+\infty} f^2 \, |M(f)|^2 \, df \ . \tag{5}$$

Analogously, measurement precision and resolution in range rate depend on the effective signal duration t_0, defined as the second moment of the squared signal envelope:

$$t_0^2 = (2\pi)^2 \int_{-\infty}^{+\infty} t^2 \, |\mu(t)|^2 \, dt \ . \tag{6}$$

The third parameter is an average measure for the internal phase structure of the signal and, as will become clear later, relates to the combined precision in range and range rate. This parameter, α, will be referred to as the effective phase constant of the signal. It is defined as

$$\alpha = 2\pi \int_{-\infty}^{+\infty} t\phi'(t) \, |\mu(t)|^2 \, dt \ , \tag{7}$$

and hence measures the linear FM content in the frequency modulation function $\phi'(t)$ of the signal. It can be shown that among all possible types of modulation and for fixed β_0 and t_0, α reaches its maximum value for linear FM.

In the presence of returns from more than a single target, the matched-filter response will be the superposition of responses to signals having different amplitudes, range delays, and Doppler shifts. Thus the tool for analyzing the resolution problem is the two-dimensional matched-filter response in delay and Doppler as used by Woodward.[1] Since it is usual practice to destroy the

fine phase information by envelope-detecting the matched-filter output, the actual quantity of interest is the complex envelope of the response and its absolute value, the real envelope. The complex envelope of the matched filter response can be written in the two alternative forms,[1]

$$\chi(\tau,\nu) = \int_{-\infty}^{+\infty} \mu(t)\,\mu^*(t-\tau)\,e^{j2\pi\nu t}\,dt \tag{8}$$

$$= \int_{-\infty}^{+\infty} M^*(f)M(f-\nu)\,e^{j2\pi f\tau}\,df\,, \tag{9}$$

where τ and ν are the differential range delay and Doppler shift, respectively, with the origin $\tau = \nu = 0$ chosen at the peak output for a perfectly matched

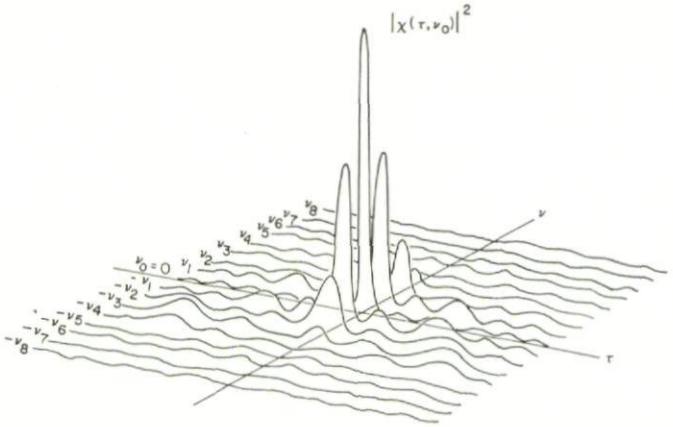


FIG. 1. *Ambiguity surface as the sum of matched-filter responses to signals mismatched in Doppler.*

filter. In other words, τ is the running time variable for the matched filter response and ν is the Doppler mismatch between return signal and "matched" filter.

The complex envelope of the matched-filter response, $\chi(\tau,\nu)$, is commonly known as the ambiguity function, alluding to the fact that subsidiary peaks of the response may, in the presence of noise, introduce an uncertainty as to the presence and location of a target. The ambiguity function is, of course, simply the filter response to a signal which is matched to the filter except for an arbitrary Doppler shift ν. The plot of $|\chi(\tau,\nu)|^2$ as a surface above the τ, ν-plane will be referred to as the ambiguity surface, and a vertical cut through this surface for constant Doppler shift ν_0 gives the power response of the matched filter for that particular value of Doppler mismatch. This permits us to indicate an ambiguity surface through a series of filter responses as shown in Fig. 1.

To summarize the important properties of the ambiguity surface, application of the Schwartz Inequality to $|\chi(\tau, \nu)|^2$ in Eq. (8) leads to

$$|\chi(\tau,\nu)| \leq \int_{-\infty}^{+\infty} |\mu(t)|^2\,dt = \chi(0,0) = 1 \tag{10}$$

In words, no subsidiary peak of the ambiguity surface can exceed the central peak in height, and this peak has a height of unity, regardless of the signal waveform. Furthermore, as already pointed out by Woodward, the total volume under the ambiguity surface equals unity,

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} |\chi(\tau, \nu)|^2 \, d\tau \, d\nu = 1 \tag{11}$$

The essence of Eqs. (10) and (11) is that, for constant signal energy, the volume under the ambiguity surface and the measurement interference it represents can be shifted around in the τ, ν-plane by changing the signal waveform, but cannot be reduced. Since changing the signal energy affects all target returns in the same manner, an increase in signal energy improves detection performance in noise but does not improve target resolution.

The volume constraint results from the definition of time and frequency as a pair of Fourier transforms, which prevents a signal from being sharply confined in both time and frequency domain simultaneously. This fact leads to the uncertainty principle of quantum mechanics and, in radar systems, to constraints on the combined resolution in range and range rate. Lacking a better terminology, we shall refer to this phenomenon as the radar uncertainty relation. For simple constant-carrier pulses, the uncertainty relation means that squeezing the matched-filter response in the delay domain to achieve better definition of the response in range must result in a corresponding widening of the response in Doppler, implying a degradation in range rate resolution. For signals of arbitrary waveforms, the effects of the uncertainty relation can be put in evidence by integrating $|\chi(\tau, \nu)|^2$ in τ and, with the use of the Parseval theorem, deriving a relation given in[3]

$$\int_{-\infty}^{+\infty} |\chi(\tau, \nu)|^2 \, d\nu = \int_{-\infty}^{+\infty} |\chi(\tau, 0)|^2 \, e^{-j2\pi\nu\tau} \, d\tau \tag{12}$$

The left integral is the area under a vertical cut through the ambiguity surface for constant Doppler shift, and as such describes the total volume distribution in the Doppler domain. The right integral is the Fourier transform of the ambiguity surface on the delay axis. Hence, Eq. (12) states that the total volume distribution of the ambiguity surface in the Doppler domain is given by the value of the ambiguity surface on the delay axis through a Fourier transform relation. This means that if the ambiguity surface is narrowed along the delay axis in order to improve measurement precision and resolution in range, the volume must spread proportionately in the Doppler domain. This is the general form of the radar uncertainty relation. For a constant-carrier pulse, spreading of the volume in the Doppler domain means widening of the central response peak and degradation of proximal target resolution. Though this need not be true for an arbitrary waveform, the spreading of the integrated volume as dictated by the uncertainty relation may lead to an equally objectionable degradation of target resolution in general. In fact, it will become clear from later discussions that changing

the signal waveform merely permits us to trade the size of the resolvable cell against the minimum strength of targets for which this degree of resolution can be achieved.

The Fourier constraint on the volume distribution must, of course, also apply in the delay domain. Interchanging the role of delay and Doppler in Eq. (12), we find the dual relation

$$\int_{-\infty}^{+\infty} |\chi(\tau, \nu)|^2 \, d\nu = \int_{-\infty}^{+\infty} |\chi(0, \nu)|^2 \, e^{2j\pi\nu\tau} \, d\nu \ , \tag{13}$$

which has analogous consequences. The total volume distribution in the delay domain depends only on the ambiguity surface on the Doppler axis.

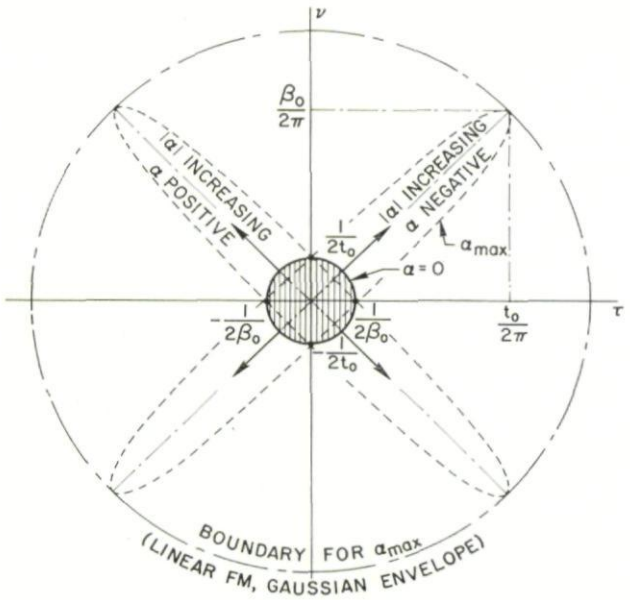


FIG. 2. *Contour of the central spike of the ambiguity surface for fixed signal band width and duration and varying phase constant.*

The minimum resolvable cell that can be provided (under circumstances to be considered later) depends on the shape of the central spike of the ambiguity surface. If this central spike is expanded into a double Taylor's series about the origin and terms of higher order than quadratic are dropped, a horizontal cut at 75 percent of peak height results in a contour of elliptical shape and widths along delay and Doppler axis that are given as the inverse effective signal bandwidth and duration, respectively. If a suitable scale is chosen for delay and Doppler axis, the situation is as shown in Fig. 2. For a given value of signal bandwidth and signal duration, the cross section of the central spike is smallest when the phase constant α is zero. This is indicated by the circle in Fig. 2. As the value of $|\alpha|$ increases, the circle is stretched into an ellipse, with the largest ellipse obtained when the signal is a linear

FM pulse of Gaussian envelope. The area of the ellipse can be calculated from the series expansion as

$$A = \pi \frac{1}{2\beta_0} \frac{1}{2t_0} \frac{1}{\sqrt{1 - (a^2/\beta_0^2 t_0^2)}} , \tag{14}$$

which shows the effect of α in increasing the area over the value for $\alpha = 0$, the circle in Fig. 2.

3. SIGNAL DESIGN FOR RESOLUTION

The constraints on the ambiguity surface of a signal, or the radar uncertainty relation, will now be used for an exposition of the general resolution problem. This approach also leads to an understanding for the significance of existing and feasible principles of high-resolution radar.

1. *Resolution as a Trade Between Target Separability and Background Visibility*

Starting with the simplest case, we assume that the radar is to resolve two point targets having the same radar cross section. The receiver response to

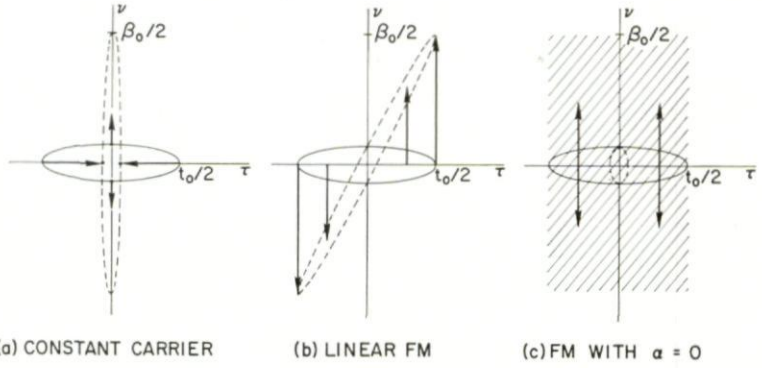


FIG. 3. *Alternative methods of reducing the width of the central spike in range delay.*

the combined return can be obtained from the superposition of the two ambiguity functions (one translated with respect to the other by the amounts of the differential range delay and Doppler shift) by taking the value of the combined function at the Doppler frequency to which the filter is matched. When the targets are separated in range only, the width of the central response peak along the delay axis must be smaller than the differential range delay if the two targets are to be resolved in the presence of interfering noise. Analogously, when the targets are separated in Doppler only, the same must apply for the central response width along the Doppler axis. The extension to the case where the targets are separated in both range and range rate is also obvious.

The effect of waveform design on the radar resolution performance is illustrated in Fig. 3. As indicated in Fig. 3a, narrowing of the ambiguity

surface along the delay axis for a constant-carrier pulse, by shortening the pulse duration, is possible only at the expense of widening the response in the Doppler domain correspondingly. If the pulse duration is held constant and the desired increase in bandwidth is achieved by modulating the carrier frequency, linear *FM* twists the ambiguity surface as shown in Fig. 3b. Though resolution in range and range rate separately now is high, it is seen that targets lying in the direction of the ridge will not be resolved. On the other hand, if the bandwidth increase is obtained by nonlinear *FM* or some other low-α modulation, in the limit with a type of modulation where the phase constant α is zero, the central peak is narrowed through smearing of the original surface in the manner of Fig. 3c.

For the case under consideration, two targets of the same cross section, a "pulse compression" signal with an ambiguity surface as shown in Fig. 3c is evidently optimum, as it will permit arbitrarily high resolution in range and range rate, provided the signal bandwidth and duration can be made sufficiently large. However, when the two targets have widely differing cross sections, the pedestal of the ambiguity surface of Fig. 3c will be objectionable. In fact, the height of the response pedestal of the strong target may be much larger than the peak of the response from the second target and thus may completely mask it. The radar then cannot "resolve" the two targets, no matter how widely separated they are in terms of the main response width. Furthermore, in the presence of a large number of targets, the clutter background introduced through the superposition of the individual pedestals may become so high that even relatively strong targets cannot be recognized.

The essence of the foregoing discussion is that target resolution cannot be specified in terms of the dimensions of the central spike of the ambiguity surface alone. Waveforms with low value of the phase constant α (and only these can truly reduce the area covered by the central spike) introduce what what may be called "self-clutter" into the radar system, and the problem of target resolution becomes one of recognizing the desired target in this clutter background, in addition to separating the returns from closely spaced targets. This clutter is the effect of the radar uncertainty relation and is basic to the problem of combined resolution in range and range rate. In the following discussion, we shall treat target resolution as the task of recognizing a particular target in the combined interference from all other targets. Since target resolvability is of interest only if even the weakest target of interest can be reliably detected in the thermal noise of the system, we shall ignore this type of noise in the consideration of resolution.

2. *Choice of Optimum Radar Waveforms*

In arriving at an optimum waveform, we have to find first the optimum ambiguity surface for a particular application and then determine the corresponding signal waveform. In a strict sense, the problem of synthesizing a waveform in accordance with a prescribed ambiguity surface has not been solved yet; however, this is of little concern for the analysis of target resolution. Once the general properties of the ambiguity surface, as outlined above, and the significance of the three signal parameters β_0, t_0, and

α are understood, it is relatively easy to visualize the interrelation between signal waveform and its associated ambiguity surface.

The returns from the illuminated targets will have delays and Doppler shifts that fall within certain boundaries in the τ, ν-plane. We shall refer to the area over which the returns are spread in this plane as the occupied target space. Since the radar uncertainty relation prevents the achievement of the ideal ambiguity surface, which is a single sharp spike, the problem of waveform design is evidently that of matching the ambiguity surface to the properties of the occupied target space. By this we mean that the volume under the ambiguity surface should be distributed over the τ, ν-plane in such a manner that it does not interfere with the measurement. To the degree to which this goal can be accomplished, waveform design will yield a improvement in the resolution performance of radar. (Since the discussion will be given in terms of delay and Doppler rather than range and range rate, it is well to point out that the target Doppler spread, and hence the size of the occupied target space, is proportional to the carrier frequency.)

As seen from Fig. 3, the problem of waveform design is simple when resolution in only one coordinate is needed. When all targets have the same range rate, we choose a constant-carrier pulse of short duration and obtain an ambiguity surface as indicated by the dashed ellipse of Fig. 3a. At some point, the signal may become so short that the transmitter cannot supply adequate energy within so short a duration. Then frequency modulation can be used to increase the signal bandwidth without decreasing the duration and the most practical type of *FM* to choose may be linear *FM* as shown in Fig. 3b. The advantage of linear *FM*, or Chirp, signals comes from the fact that the signal stays matched to the filter (or is properly "compressed") even if it has the wrong Doppler shift. Furthermore, a Chirp signal is relatively easy to generate and process. The practical problem is to design the signal such that the residual volume in the vicinity of the delay axis is small or, in more practical terms, that the remaining range "side lobes" after compression are very much smaller than the central peak. Theoretically, however, any other type of *FM* would serve just as well, and there are no limitations on achievable resolution.

The situation is similar when all targets have the same range and only range rate resolution is needed. Since high range rate resolution requires long signals, the limitation is not the transmitter power but the receiver isolation. For long-range operation, the signal duration cannot exceed the width of the occupied target space in the delay domain, so that coherent pulse trains must be used when still higher range rate resolution is required than corresponds to the limiting value of the single-pulse duration. The properties of pulse train type signals will be considered later.

Of more interest is waveform design when targets have to be resolved in both range and range rate. In a common practical situation, we know no more about the properties of the target environment than that targets may have a certain spread in range delay and Doppler. In other words, the only prior information available and usable is that the targets fall within certain boundaries in target space. The optimum ambiguity surface for this case

consists of a narrow central spike, with the remainder of the volume spread out into a uniformly low pedestal. This ambiguity surface is usually referred to as a "thumbtack" surface, and its characteristics may be deduced from the Fourier constraints of Eqs. (12) and (13): for a central spike width in the order of $1/\beta_0$ and $1/t_0$ along delay and Doppler axes, respectively, essentially all of the unity volume must be contained within $|\tau| \leq t_0$ and $|\nu| \leq \beta_0$. Furthermore, for the high time-bandwidth product necessary to achieve a thumbtrack type surface the volume in the spike is evidently negligible, and the height of the pedestal must be $1/\beta_0 t_0$ to satisfy the volume constraint. These over-all dimensions of the thumbtack surface are indicated in Fig. 4.

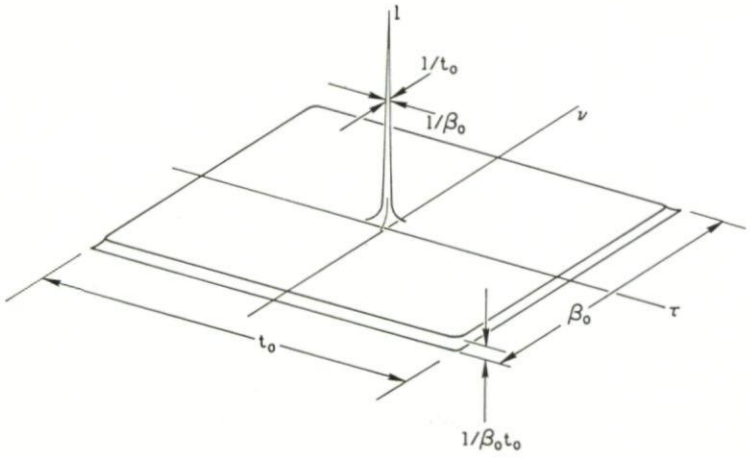


FIG. 4. *Thumbtack ambiguity surface and its constraints.*

As will be shown later, the pedestal of the thumbtack surface can lead to a serious degradation of resolution performance even though it is relatively low for high time-bandwidth product signals. In those cases where the occupied target space has a small size, it may be preferable to use an ambiguity surface where the volume from around the central spike is shifted to other places in the τ, ν-plane. No interfering self-clutter will then be generated as long as the cleared area is at least twice as wide as the occupied target space in both delay and Doppler domain. This raises the question as to how large an area can be cleared.

Along either axis, the ambiguity surface is determined by the autocorrelation function of the signal in the respective domain. Since only signals having pulse train structure can have an autocorrelation function that is zero over an extended interval, it is clear that the desired ambiguity surface can be achieved only with pulse trains in the time domain or "pulse trains" in the frequency domain. Choosing the former for purposes of illustration, uniform repetition of an arbitrary signal at a repetition period T will deposit "volume" at periodic intervals on the delay axis, as indicated by the marks in Fig. 5. Because of Eq. (12), most of the volume in the τ, ν-plane must then be concentrated in strips parallel to the delay axis and

repeated at intervals $1./T$. If, in this process, volume is deposited on the Doppler axis, it must fall within the horizontal strips, and through Eq. (13) then confines the volume to the vertical strips in Fig. 5. The result is the well-known ambiguity surface of the uniform pulse train, with spikes of unity height at the intersecting points of the grid. As shown by the rectangle in Fig. 5, the maximum size of the occupied target space which the pulse train can accommodate without introducing serious self-clutter is unity.

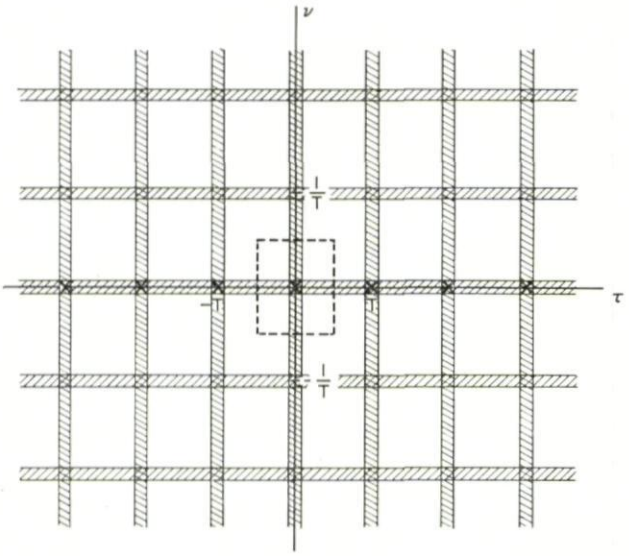


FIG. 5. *The "clear area" of signals with the structure of pulse trains.*

It can be shown that the size of the clear area cannot be increased through either intra-pulse or inter-pulse modulation of the pulse train. As is seen from Fig. 3c, modulation with small phase constant α achieves narrowing of the response peak through smearing of the undesired volume into a pedestal, thus introducing self-clutter. High-α modulation translates ridges or spikes in the manner of Fig. 3b and merely distorts the shape of the clear area without altering its size. This behavior of pulse trains is studied in more detail in another paper.[6]

It should be mentioned that the "clear area" cannot be strictly free from all volume. As a consequence of the definition of time and frequency as a pair of Fourier transforms, no signal can have extended zero intervals in both time and frequency domain. Thus there will always be some volume left between the individual spikes of the ambiguity surface; for example, a pulse train in time will have some volume spread out within the vertical strips of Fig. 5. In some cases, the interference from such secondary response lobes can be very objectionable, yet these effects are second order compared to those from the full pedestal.

3. *Approximation of the Thumbtack Surface Through Practical Waveforms*

In the preceding section, we have found that when resolution in only one coordinate, range or range rate, is needed, the signal design is dictated only by practical considerations. When the occupied target space is smaller than unity, the proper choice for the signal is the pulse train, where this term includes a signal consisting of the periodic repetition of frequency bands. In the general case where the illuminated target space is large compared to

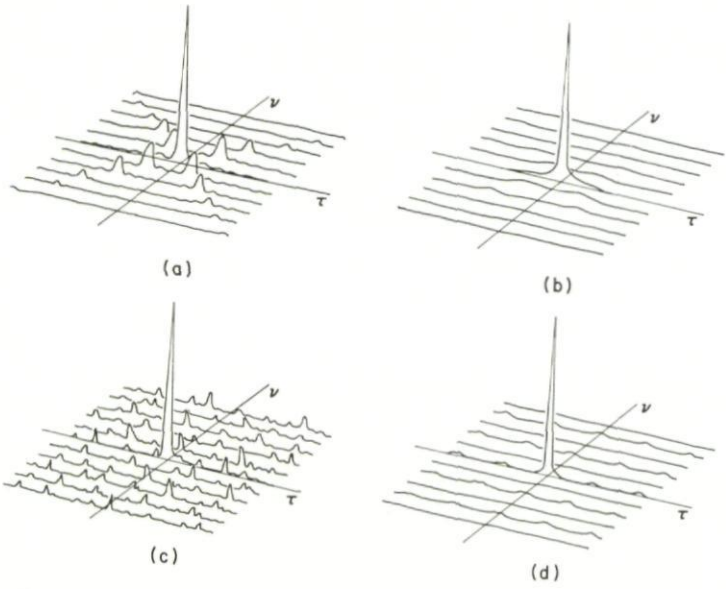


FIG. 6. *Representative examples for practical approximation of the thumbtack surface. (a) V-type Chirp, (b) Quadratic FM, (c) Phase-reversal coding, (d) Pulse train with pulse-to-pulse coding.*

unity, however, the radar should be using a signal having a thumbtack ambiguity surface. We shall now add some remarks about waveforms which approximate this surface.

It is clear from Fig. 2 that the signal must have a large time-bandwidth product and a phase constant α which approaches zero. As it is not practical in high-power radar to use signals of other than roughly rectangular envelope the desired increase in the signal bandwidth can be achieved only by *FM* or *PM* of the carrier. From the definition of the effective phase constant, Eq. (7), we then find that α is zero whenever the mean of the *FM* function $\phi'(t)$ is zero. One way to achieve this is to use symmetrical *FM*, such as the combination of the frequency up-sweep and down-sweep of the simple Chirp signal, or "*V*-type" Chirp. The corresponding ambiguity surface is approximately indicated by the contours for fixed Doppler shift in Fig. 6a. Since each half of the signal is a Chirp signal and contains a strong linear

FM component, the ambiguity surface still shows the pronounced ridges reminiscent of a Chirp signal. For a further suppression of these ridges, we should use symmetrical *FM* without linear segments, for example, quadratic *FM*. The corresponding ambiguity surface is sketched in Fig. 6b.

With signals having smooth modulation functions, it is unavoidable that much of the volume is concentrated in the central portion of the ambiguity surface, giving rise to near sidelobes of appreciable magnitude. A more even distribution of the volume can be achieved with modulation functions of random character, such as noise modulation or phase reversal coding in accordance with some pseudo-random binary code. The ambiguity surface then takes the form of Fig. 6c, with a fairly uniform distribution of the volume but occasional subsidiary peaks of appreciable magnitude. The practical difficulty consists in selecting the modulation code such that the interfering spikes stay below a given level.

When the desired range rate resolution is so high that the signal duration exceeds the target spread in range delay, a coherent pulse train must be used. However, the pulse train must employ some kind of pulse-to-pulse coding to smear the ambiguous spikes of the uncoded pulse train into a low pedestal.[6] Jittering of the repetition interval, changing the waveform from pulse to pulse, or pseudo-random frequency hopping are some codes which produce the desired smearing effect. Codes with high phase constant α, such as linear frequency hopping from pulse to pulse, must of course be ruled out since they result in a translation of the volume without any leveling of the subsidiary spikes.

4. FUNDAMENTAL LIMITATIONS ON RADAR RESOLUTION

It was found earlier that in the attempt to decrease the width of the central spike of the ambiguity surface and improve proximal target resolution, a background of self-clutter is generated which in turn may prevent the recognition of targets altogether. We shall now demonstrate that this self-clutter imposes a fundamental limit on the resolution performance of radar.

Assume that the targets illuminated by the radar beam are constrained to an interval $\Delta\tau$ in range delay and an interval $\Delta\nu$ in Doppler, and the size of this (occupied) target space, $\Delta\tau\Delta\nu$, is large compared to unity. The optimum ambiguity surface under these conditions is the thumbtack surface. When many targets are present, and this is where the limitations on resolution become apparent, the superposition of the many pedestals from returns having different delays and Doppler shifts will have an averaging effect on the combined pedestal so that its detail structure is of little interest. Hence we can estimate radar resolution performance on the basis of an ideal thumbtack surface and, for the present purpose, ignore the problem of approximating the ideal thembtack surface by a practical waveform.

The relation between the size of the occupied target space and the spread of the pedestal is depicted in Fig. 7. As the first case, we assume that the

target space contains N discrete point scatterers of a total cross section $\delta_0 = \sigma_t/N$. Furthermore, N will be assumed large enough to allow treating the superposition of the pedestals as a power superposition, that is, superposing the pedestals of the ambiguity *surfaces* rather than ambiguity *functions*. (This actually corresponds to the calculation of the variance of the combined clutter which, in the case of many contributions, will very nearly halve a Gaussian distribution, so that target recognition in the clutter becomes the well-known problem of signal detection in Gaussian noise.)

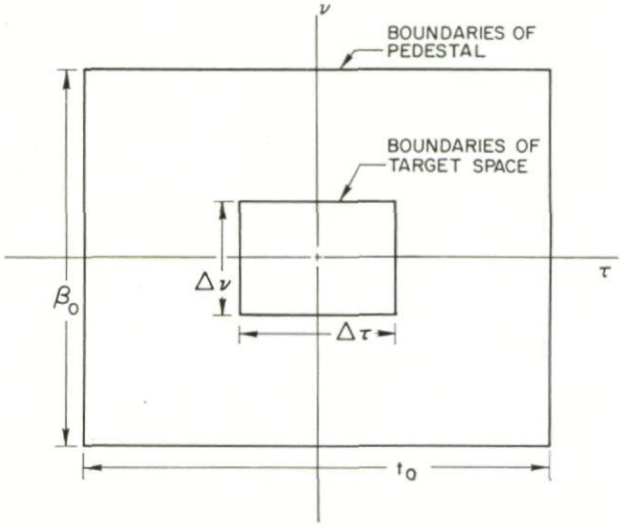


FIG. 7. *Occupied target space and spread of the pedestal of the ambiguity surface.*

Since the peak of the normalized thumbtack surface has a height of unity and a pedestal of height $1/\beta_0 t_0$, proper amplitude scaling of each return signal requires that the heights of both peak and pedestal be multiplied by the radar cross section of the corresponding target. Superposing the scaled ambiguity surfaces from all N targets, the height of the combined pedestal becomes

$$H_p = \sum_1^N \frac{\sigma_\nu}{\beta_0 t_0} = \frac{\sigma_t}{\beta_0 t_0} = N \frac{\sigma_0}{\beta_0 t_0} \ . \tag{15}$$

For a target of cross section σ_ν, the signal-to-clutter ratio thus is

$$\frac{S_\nu}{C} = \frac{\sigma_\nu}{Hp} = \beta_0 t_0 \frac{1}{N} \frac{\sigma_\nu}{\sigma_0} \ , \tag{16}$$

and for a target of average cross section,

$$\frac{S_0}{C} = \frac{\beta_0 t_0}{N} \ . \tag{17}$$

The value of the signal-to-clutter ratio needed for proper target identification in the background will depend on the requirements of the particular

application. In most instances, a minimum signal-to-clutter ratio in the order of 6 db might be typical. However, in order to avoid obscuring the essence of the discussion by numerical factors, we shall here assume a minimum required signal-to-clutter ratio of unity. Equation (17) then shows that the maximum number of targets which can be accommodated if a target of average cross section is to be resolved is given by the time-bandwidth product of the signal,

$$N_{\max} = \beta_0 t_0 \ . \tag{18}$$

Provided the area of the pedestal is larger than the occupied target space, as in Fig. 7, the target capacity of the radar can be increased by increasing the time-bandwidth product of the signal. It is seen from Eq. (16) that if a target of lower than average cross section is to be resolved, the target capacity of the radar decreases accordingly. Specifically, a target m db below average can be resolved only if the number N of targets is m db below $N_{\max}$ as given by Eq. (18). As an example, if 100 targets are present and targets with a cross section 20 db below average are to be resolved, the transmitted signal must have a time-bandwidth product of 10,000, a rather expensive requirement for a practical radar.

In some applications, the relation between size of the occupied target space and area of the pedestal may be the reverse from that of Fig. 7, and it is of interest to consider resolution when $\Delta\tau\Delta\nu$ is large compared to $\beta_0 t_0$. Now it is preferable to introduce the average spacing of the targets in delay and Doppler, ξ_τ and ξ_ν, rather than the total number of targets in target space. The number of contributors to the combined pedestal is given as the number of targets within the area of the single pedestal, or

$$N = \frac{\beta_0 t_0}{\xi_\tau \xi_\nu} \ , \tag{19}$$

and the signal-to-clutter ratio for a target of cross section σ_ν is

$$\frac{S_\nu}{C} = \xi_\tau \xi_\nu \frac{\sigma_\nu}{\sigma_0} \ . \tag{20}$$

As shown by the preceding result, target resolvability in the clutter now depends only on the target strength and the average target density, but not on the time-bandwidth product of the signal. Though increasing the time-bandwidth may improve proximal target resolution for targets of sufficient strength, it does not improve resolvability in the self-clutter. For a target of average cross section, setting the signal-to-clutter ratio in Eq. (20) equal to unity yields

$$\xi_\tau \xi_\nu = 1 \tag{21}$$

This means that a target of average cross section can be resolved only if the target density is no higher than one target per unit area in target space. In other words, the size of the resolvable cell cannot be made smaller than unity, which is the resolvable cell that can be provided by a simple constant-carrier pulse. If the target density is higher than one target per unit area,

only targets proportionately stronger than the average target can be resolved, with a strict correspondence between the increase in target density and increase in the strength of resolvable targets.

The situation is similar for continuously distributed targets. We analyze this case by assuming that the target cross section in each cell of dimensions $1/\beta_0 \times 1/t_0$, the area of the central spike of the ambiguity surface, is concentrated into a single point target. The task of the radar then is to measure the cross section of each such target and obtain a "map" of the cross section distribution in target space. Of course, the self-clutter will interfere with the measurement, and the cell size will have to be large enough to ensure that the cross section per cell, or the equivalent point target, is so large that it can be recognized in the clutter. We again have the trade-off between cell size and adequate target strength.

When the occupied target space is large compared to unity, the thumbtack ambiguity surface is again optimum. The relation between target space and pedestal is as shown in Fig. 7, and the number of scatterers contributing to the clutter background equals the number of cells of size $1/\beta_0 t_0$ within the occupied target space, or

$$N = \frac{\Delta_\tau \Delta_\nu}{1/\beta_0 t_0} \tag{22}$$

The signal-to-clutter ratio for a cell containing the target cross section δ_ν is, from Eq. (15),

$$\frac{S_\nu}{C} = \frac{1}{\Delta\tau\Delta\nu}\frac{\sigma_\nu}{\sigma_0} = \frac{1}{\Delta\tau\Delta\nu}\frac{\sigma_\nu^*}{\sigma_0^*}\ , \tag{23}$$

where the asterisk is used to designate cross section densities in target space. It is seen that an average cross section density can be properly mapped only if the total occupied target space is unity in size or smaller. At those portions in target space where the cross section level is m db below average, correct mapping (or target "resolution") is possible only when the occupied target space has a size m db below unity. Conversely, for target spaces larger than unity, mapping of the distribution of cross section is feasible only where the cross section density is proportionately higher than average. For a given target space, changing the time-bandwidth product of the signal thus merely permits trading the size of the resolution cell against the cross section level for which a resolvable cell of size $1/\beta_0 t_0$ can be truly provided in the self-clutter.

Considering the other case of interest, a target space large compared to the area of the pedestal, the number of cells contributing to the background of self-clutter is

$$N = \frac{\beta_0 t_0}{1/\beta_0 t_0} = (\beta_0 t_0)^2, \tag{24}$$

and the signal-to-clutter ratio for a cell with cross section σ_ν is

$$\frac{S_\nu}{C} = \frac{1}{\beta_0 t_0}\frac{\sigma_\nu}{\sigma_0} = \frac{1}{\beta_0 t_0}\frac{\sigma_\nu^*}{\sigma_0^*} \tag{25}$$

A resolution commensurate with the area of the central spike of the ambiguity surface, or a cell of dimensions $1/\beta_0 \times 1/t_0$, can evidently only be achieved in those portions of target space where the cross section density is higher than average by a factor $\beta_0 t_0$ or more. Where the cross section density is of average strength, the truly resolvable cell is unity in size, and when the cross section density is still lower, the size of the resolvable cell is proportionately larger than unity.

The preceding results have revealed very fundamental limitations on the combined range and range rate resolution achievable with radar when the occupied target space is large compared to unity. When it is equal to unity or smaller, these limitations can be avoided by using pulse train type signals to clear the area around the central spike of the ambiguity surface and achieve an arbitrarily high resolution, except for the effects of the residual volume in the "clear" area. As the occupied target space exceeds unity by a larger and larger factor, the benefits from being able to clear an area of unity size will become of less value. Furthermore, we shall have to smear the ambiguous spikes of the ambiguity surface of the uniform pulse train by pulse-to-pulse coding, and in this process approximate the thumbtack surface. Hence, this type of ambiguity surface must be used whenever the size of the occupied target space significantly exceeds unity, and we than have the limitations on resolution as discussed above. Of course, in those simple cases where resolution in only one coordinate is needed, range or range rate, the limitations on achievable resolution are only of practical nature, such as the difficulties of realizing a large time-bandwidth product or reducing the spurious responses to the desired low level.

5. PRACTICAL IMPLICATIONS OF THE CONSTRAINTS ON RESOLUTION

It is instructive to illustrate the practical significance of the preceding results by considering two typical applications for high-resolution radar. As

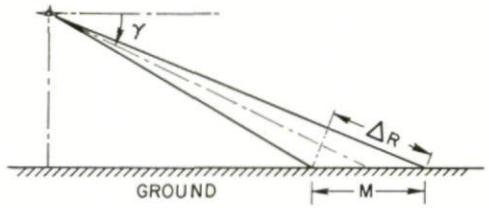


FIG. 8. *Geometry for synthetic radar operation.*

the first example, we choose the synthetic aperture ground mapping radar, a system whose task it is to provide a map of the continuous cross section distribution on the ground. As a second example, typical for resolution of high-density discrete targets, we consider a radar operating against dense-target clouds such as aircraft or missiles accompanied by chaff and decoys.

The geometry of synthetic aperture radar is shown in Fig. 8. Briefly summarizing its operation, as described by Cutrona,[7] the radar flies along a

straight path with velocity v and periodically illuminates a ground swath of width M. The return signals are stored and then processed such as if they had been received by a long array antenna with a correspondingly narrow beam in azimuth. As an alternative and, for our purposes, more interesting interpretation, the radar uses coherent processing of a very long signal to achieve high resolution in Doppler which, because of the known relative motion of radar and target, can be converted into high resolution in azimuth. Range resolution is obtained as in any conventional high-resolution radar.

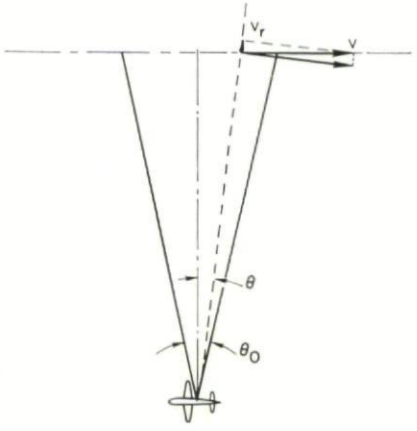


FIG. 9. *Derivation of the target Doppler spread.*

Referring to previous discussions, in a ground mapping radar the targets are continuously distributed over large intervals in range and Doppler. The occupied target space in the sense defined above is that portion of the total target space which is actually illuminated by the radar beam. Having found that proper mapping is achievable only when the occupied target space does not exceed unity, we can immediately conclude that the antenna aperture must be large enough to limit the size of the illuminated target space to unity, in which case the interfering clutter background is avoided. Then there will be no *theoretical* limit on achievable resolution, but the constraint on the size of the occupied target space now imposes a limit on the size of the map that can be obtained.

Converting the limit on the occupied target space into a limit on map size, the spread of the targets in range delay is found from Fig. 8 as

$$\Delta\tau = \frac{2\Delta R}{c} = \frac{2M\cos\gamma}{c} \quad . \tag{26}$$

The target spread in Doppler can be calculated from the relative target motion as indicated in Fig. 9. For a target at angular position θ, the Doppler shift is

$$\nu = 2\frac{v_r}{\lambda} = 2\frac{v}{\lambda}\sin\theta \approx 2\frac{v}{\lambda}\theta \quad , \tag{27}$$

where the approximation is good because the antenna beam is rather narrow in practice. (Though target acceleration is not negligible in synthetic aperture applications, it is predictable from target range and hence can be taken out in data processing.) The total Doppler spread of the targets within the beam at any one time is determined by the azimuth width θ_0 of the beam as (Eq. [27])

$$\Delta\nu = 2\frac{v}{\lambda}\theta_0 \ , \tag{28}$$

and the unity constraint on the occupied target space becomes

$$\Delta\tau\Delta\nu = 2\frac{\nu}{\lambda}\theta_0\frac{2M\cos\gamma}{c} = 1 \ . \tag{29}$$

From Eq. (27), the relation between resolution in Doppler and resolution in azimuth angle is

$$\delta\nu = \frac{2\nu}{\lambda}\delta\theta \ . \tag{30}$$

We now replace the angular resolution $\delta\theta$ by the lateral resolution in azimuth at target range R, $a = R\delta\theta$. Furthermore, Doppler resolution is in the order of the inverse signal duration, and this duration is given by the time it takes the target to traverse the antenna beam; hence

$$\delta\nu = \frac{1}{T} = \frac{2v}{\lambda}\frac{a}{R} \ , \tag{31}$$

and with $T = \theta_0 R/v$, Eq. (31) and the constraint of Eq. (29) lead to

$$M = \frac{1}{2\cos\gamma}a\frac{c}{v} \ . \tag{32}$$

This result states that the self-clutter in high-resolution radar limits the achievable swath width. For a given vehicle velocity v, the swath width decreases proportionately to the width of the azimuth resolution element, a. As an ancillary result, for a given resolution in azimuth, the achievable swath width is inversely proportional to the velocity of the vehicle carrying the radar, which means that low-velocity vehicles can map wider swaths than high-velocity vehicles. This is an expression of the fact that the size of the occupied target space in range delay and Doppler depends on the vehicle velocity through the relation between Doppler frequency and velocity as given in Eq. (27).

For the second example, a radar operating against a target "cloud", the resolution performance again depends on the size of the target space illuminated by the beam. When the occupied target space is unity in size or smaller, coherent pulse trains can be used to produce an ambiguity surface with a clear area around the central spike, and resolution is limited only by the residual volume in the clear area. In practice, this leads to a limitation on the size of weak targets that can be seen in the presence of strong targets and becomes a matter of successful system design. On the other hand, when

the occupied target space is large compared to unity (and this can be made deliberately so in weapon system design), the problem is much more fundamental. This case will be considered below.

For large target spaces, the optimum ambiguity surface is the thumbtack surface, and for high resolution requirements the sketch of Fig. 7 applies. Using Eq. (16), the signal-to-clutter ratio for a target of cross section σ_v can be written

$$\frac{S_v}{C} = \beta_0 t_0 \frac{\sigma_v}{\sigma_t} \,. \tag{33}$$

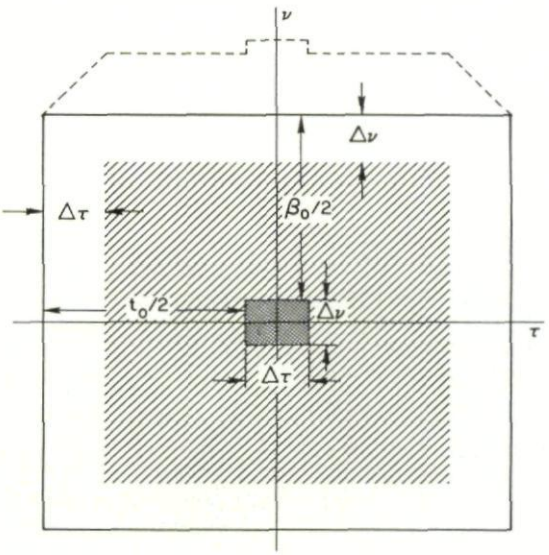


FIG. 10. *Spreading of the clutter in target space for "high-resolution" signals.*

where σ_t is the total target cross section within the target space illuminated by the radar beam. In order for a target to be recognizable in the clutter background, its cross section must evidently be at least as large as the fraction $1/\beta_0 t_0$ of the total cross section σ_t. Although weak target resolution can, in theory, be improved by choosing higher time-bandwidth products, there exist definite practical limitations on how a time-bandwidth product an operational radar can employ. Furthermore, after a certain point, the decrease in the average pedestal height of the ambiguity surface for larger time-bandwidth product becomes meaningless, as it will be progressively more difficult to keep the subsidiary peaks above the average pedestal level low enough to prevent weak target masking by such individual peaks rather than the average clutter level.

The difficulties of dense-target resolution are even more serious than suggested above. The lowering of the clutter level by increasing the time-bandwidth product of the signal is, of course, paid for by a corresponding spreading of the pedestal, as indicated in Fig. 10. This shows that the

signal-to-clutter ratio of Eq. (33) exists not only for targets within the densely occupied target space, but over an area which can be very much larger. In other words, if the signal-to-clutter ratio for a target within the target cloud is insufficient for proper target resolution, it will remain insufficient even if the target moves outside the cloud. Practically, the target of interest may have appreciably different range and range rate than the bulk of the targets and still be effectively hidden. The penalty paid for decreasing the size of the resolution cell is a corresponding spreading of the masking effects of the target cloud beyond its boundaries in range and range rate. Again, these effects trace back to the radar uncertainty relation, and their severity can be easily appreciated by substituting practical numbers in the relations derived above.

6. CONCLUDING REMARKS

Modern radar offers the possibility of significantly reducing the size of the resolvable cell in target space from the value of unity achieved with constant-carrier, single-pulse signals. However, aside from the practical difficulties of implementing such a radar, increasing the time-bandwidth product of the transmitted signal introduces self-clutter which itself limits radar resolution performance. It was shown that, as a consequence of the radar uncertainty relation, arbitrarily high resolution can even in theory be achieved only when the total occupied target space does not exceed unity in size, in which case pulse train type signals must be used to generate an ambiguity surface with a clear area around the central spike. For large occupied target spaces, and given characteristics of the target environment, signal design merely permits trading the size of the resolvable cell against the target strength for which this resolvable cell can actually be provided. Two practical examples were given in order to illustrate that these imitations are not merely of theoretical interest but do lead to very practical constraints on the resolution performance of even the most sophisticated radar.

REFERENCES

1. Woodward, P. M., *Probability and information theory, with applications to radar*, Pergamon Press, New York, 1953.
2. Siebert, W. M., A radar detection philosophy, *Trans. IRE* **IT-2,** 204, 1956.
3. Westerfield, E. C., Prager, R. H. and Stewart, J. L., Processing gains against reverberation (clutter) using matched filters, *Trans. IRE* **IT-6,** 342, 1960.
4. Fowle, E. N., Kelly, E. J. and Sheehan, J. A., Radar system performance in a dense-target environment, *IRE Internat. Conv. Rec.*, **4,** 136, 1961.
5. Gabor, D., Theory of communication, *J. Inst. Electr. Engrs.*, **93,** Part III, 429, 1946.
6. Rihaczek, A. W., Radar resolution properties of pulse trains, *Proc. Inst. Electron. Elect. Engrs.* **52,** 153, 1964.
7. Cutrona, L. J., Vivian, W. E., Leith, E. N., and Hall, G. O., A high-resolution radar combat-surveillance system, *Trans. Inst. Radio Engrs.*, **MIL-5,** 127, 1961.

CHAPTER 7

MEASUREMENTS OF THE REFLECTION FROM THE GROUND OF S-BAND SIGNALS

A. R. DOMVILLE
G.E.C., Stanmore, Middlesex, England

Measurements of the reflection of S-band pulse signals from various types of ground have been made at vertical incidence: the measurements were made in conditions where the return signal was beam-width-limited.

The mean reflection from most ground surfaces was found to be between 5 *and* 15 *db below, and from water to be about* 1 *db above that from a surface scattering uniformly without loss. Most surfaces reflected primarily by scattering, but a few gave predominantly specular reflection, and in this case the amplitude of the return signal depended strongly on the shape and inclination of the surface; the extreme values of reflection were* +7 *and* −31 *db with respect to a uniform scattering surface.*

The standard deviation and amplitude distribution of most surfaces is nearly Rayleigh, but significant differences have been found for near specular reflectors. Over undulating country, extended and multiple peaked echoes have been observed.

1. INTRODUCTION

This paper describes and gives the results of a series of experiments carried out to measure the reflection of S-Band radio transmissions from various types of ground surface. The measurements were made at near vertical incidence with pulse signals.

The terrains over which measurements were made were mostly chosen for their homogeneity and flatness. The sea was taken as a convenient reference surface which could be relied upon to give consistent results as a check on equipment performance, although it was not used for calibration.

A few measurements were made over farmland and urban areas which were expected to provide intermediate amplitudes of return signal, but particular attention was given to terrains such as forest and dry sand expected to provide low amplitude echoes.

In addition to these homogeneous and flat terrains, the effects of undulating country and of hillsides were also studied.

The return signal from a transmitter in a moving aircraft will show a pulse-to-pulse variation of amplitude for any terrain which is not absolutely flat and homogeneous. Therefore not only the mean value for the reflected signal, but also the way it varies is indicated. Distortion of the echo pulse

shape such as stretching or multiple peaking may affect the operation and accuracy of any altimeter device, and measurements were therefore made of such distortions.

2. EXPERIMENTAL DETAILS

The measurements were carried out with a transmitter and receiver installed in a small aircraft, and arranged to measure the attenuation between the transmitter and received signals.

The following characteristics of the radar system used are relevant to interpretation of the results:—

Frequency	S-Band
Transmitter	Pulses of approximately 0.2 μs duration, at 1000 p.p.s.
Aerial	Approximately 24° over the −3 db points with 16 db on-axis gain. Aerial aligned vertically downwards.
Calibration	By cable of known attenuation, correction being made for cable temperature.

The aircraft was normally flown between 1000 and 3000 feet above the terrain being studied, at an air speed between 150 and 180 knots. Recordings were generally made in bursts of about 10 seconds, and the following photographic records were taken:—

(1) The peak amplitude of each individual return pulse was presented on a C.R.T. and recorded on continuously moving film.
(2) Photographs of an 'A' scope were taken at one per second.
(3) The principal equipment variables were presented on meters and the meter bank was photographed once a second.
(4) The ground was photographed by a mapping camera.

3. EXPERIMENTAL ACCURACY

The consistency of results was better than 0.5 db, and the absolute accuracy is thought to be within ± 1.5 db.

4. REFLECTIVITY DATUM

The measured amplitudes of the return signals are here expressed in terms of the power loss or gain relative to an arbitrary datum derived from a simple model of the ground. The simple model selected is that in which the incident power is reflected by scattering isotropically over the upper hemisphere without loss by absorption or change of polarisation. Results obtained in this way are often called values of "$f(\theta)$". Theory shows that such a scattering surface gives a ratio of received to transmitted power:—

$$\left(\frac{P_r}{P_t}\right)_{\text{scat}} = \frac{\lambda^2 G}{16\pi^2 h^2}$$

where λ is the wavelength, G is the aerial gain, h is the height of the aerial from the ground. The equipment was used in such a way as to give measured values of the ratio of received to transmitted power, and the reflection from any terrain can then be quoted in terms of the relation between the measured and the model values, i.e.

$$\text{reflection gain} = \left(\frac{P_r}{P_t}\right)_{\text{meas}} \cdot \frac{16\pi^2 h^2}{\lambda^2 G}$$

Had Lambert's Law been taken as the basis of the model, reflection gains 3 db lower than those quoted here would have been obtained.

It is also interesting to consider the case of specular reflection rather than scattering. For specular reflection with a horizontal surface, a vertical beam and no reflection loss, it can be shown that:—

$$\left(\frac{P_r}{P_t}\right)_{\text{spec}} = \frac{\lambda^2 G^2}{64\pi^2 h^2}$$

With the value of gain, G, used in these experiments, this would give a reflection 12 db greater than the scattering model. If the specular surface is not horizontal, the return signal may be very small, being derived only from the off axis region of the aerial pattern.

5. ANALYSIS OF RESULTS

1. *Sampling of Data*

The main work of analysis has been concerned with the continuous film record of echo amplitudes, illustrated in Fig. 5. From each recording burst of about 10 seconds, a block of 500 or 1000 pulses was selected for analysis, usually by reference to the mapping camera record to ensure that the area chosen was homogenous. For terrains which could be considered reasonably uniform, this was the only method of selection used, but for snow covered forests and desert sand where the echo amplitude varies widely, a number of additional blocks were selected to ensure that the full range of reflectivity was included.

The length of the block, 500 or 1000 pulses, was chosen to reduce sampling errors to negligible proportions.

2. *Reflection Gain*

The values of mean reflection gain, $f(\theta)$, quoted were obtained by taking an average over the peak amplitudes of the pulses in a block, and then expressing it in a decibel form. Where a number of measurements of one terrain type were available, they were themselves averaged (in decibels) to give an average for that terrain type. The standard deviation of the block averages has also been calculated to give a measure of the uniformity of a particular terrain type.

3. *Amplitude Distribution*

For one block from each terrain type an amplitude distribution of the peak echoes within the block was obtained (Figs 1 and 2).

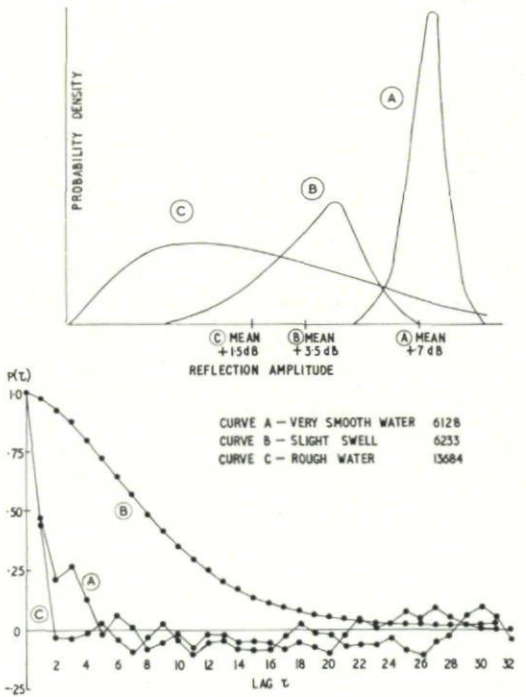


FIG. 1. *Distributions and autocorrelations, water, with various surface states.*

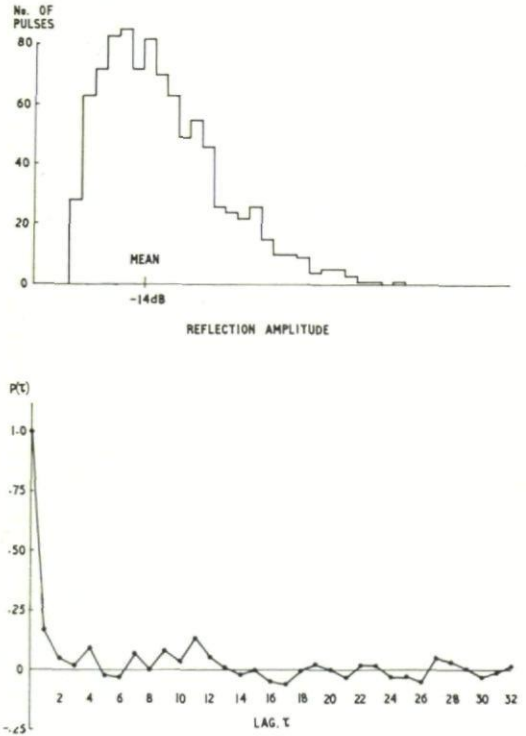


FIG. 2. *Distribution and autocorrelation, forest* (9074).

For terrains such as forest or rough sea which reflected predominantly by scattering the distribution was closely Rayleigh, but where a specular component was present the amplitude distribution was narrower.

In addition to plotting the distribution, the standard deviation of echo amplitude within a block has been calculated and expressed as a proportion of block average. Also the fading range, that is the range of echo amplitude between the level below which only 5% of the returns are found and the level below which 95% are found has also been quoted.

The distribution takes no account of the order in which echo amplitudes occur. To describe this aspect of echo amplitude variation the autocorrelation function has been plotted (Figs. 1 and 2).

4. *Tables of Results*

Table 1 summarises the results under terrain types.

Table 2 is a list of all the results obtained with a very brief description of each terrain. Detailed descriptions are available of the terrain samples but are not included here.

5. *'A'-Scope Records*

Photographs of the shape of the returned echo pulse are shown in Figs. 3, and 5. Figures 3a,b,c,d, are multiple exposures of several returns on the same film, but in the case of Fig. 5 the film was moved vertically from echo-to-echo to show the variation of pulse shape. The timing wave included is at 4 Mc/s.

6. DISCUSSION OF RESULTS

6. *Water*

The results for water have been divided into three groups. The first group is represented by one measurement made over a lake where the water was very still, and an almost pure specular reflection was obtained. The 'A'-scope photograph, Fig. 3a is an almost perfect reproduction of the transmitter pulse.

The second group is for sea states between 0 and 1. The reflection gain, +1.5 db, is greater than that expected from a scattering surface the distribution is less than Rayleigh, and the 'A'-scope photograph (Fig. 3b) shows a variation in echo amplitude. The echo is therefore partly specular and partly scattering.

The third group is for sea states greater than 1, and here the results show great consistency. The standard deviation of the block averages is only 0.5 db The distribution of echo amplitude within the block is closely Rayleigh and the 'A'-scope photograph (Fig. 3c) shows a wide variation of amplitude.

'A'-scope photographs for water show the echo at constant range which appears to be a characteristic of this terrain type.

7. *Forest*

All the records taken of reflection from forest areas show the characteristics of scattering surfaces. The 'A'-scope photograph (Fig. 3d) shows variation in range which is characteristic of returns from land.

TABLE 1
SUMMARY OF RESULTS BY TERRAIN TYPES

Terrain type	Water			Forest			Plain						
				Coniferous									
	Very smooth	Sea State 0 to 1	Sea State >1	Dry	Snow-covered	Deciduous	Snow-field	Farm land	Marsh	Desert sand	Semi-desert	Sand-dunes	Urban
Mean reflection gain, db	+7	+2	+1.5	−13	−12.5	−15	−9.5	−11	−11.5	−5	−5.5	−10.5	−5.5
No. of block averages used	1	11	6	3	13	5	10	3	3	9	2	17	7
S.D. of block averages, db	—	2	0.5	1	2	1	2	3.5	1.5	2.5	—	5	1.5
Extreme block averages by inspection, db					−9/−16		−4.5/ −16.5			−1/ −14		−4/−31	
Ratio of s.d. to mean amplitude within typical block	0.06	0.19	0.56	0.54	0.55	0.6	0.47	0.56	0.56	—	0.5	—	0.6
Fading Range/5% to 95%/of amplitudes within typical block, db	2	6	18	18	14	—	13	15.5	16.5	—	—	—	15

TABLE 2
INDIVIDUAL RESULTS

Reference No.	Height feet	Terrain	Mean reflection gain, db
2693	2000	Sea, state 3 to 4	−1
2724	1500	Sea, state 3 to 4	−½
2834	3000	Snow	−8½
2855	3000	Sea	+2½
2869	2000	Sea	+1
2892	1000	Sea	−2
2907	1500	Sea	+1
3069	3000	Sea, state 1	+2
3090	2000	Sea, state 1	+1½
3140	1000	Sea, state 1	+½
3259	1500	Snow	−9½
3315	2000	Snow	−13½
3350	3000	Snow	−11½
3538	2000	Forest and snow	−14
3576	1500	Forest and snow	−13½
3648	2000	Conifers and snow	−14
3692	1500	Conifers and snow	−15½
3737	800	Conifers and snow	−15½
3875	800	Conifers and snow	−18
4091	800	Conifers and snow	−12½
4171	2000	Sea, state 1	+1
4192	1500	Sea, state 1	+½
4210	800	Sea, state 1	+1½
6003	2000	Conifers and snow	−16
6077	1000	Conifers and snow	−16
6128	2000	Smooth water	+7
6191	1000	Sea, state ½	−½
6206	1500	Sea, state ½	+2
6233	2000	Sea, state ½	+3½
6312	2000	Snow field	−11
6338	1500	Snow field	−13½
6386	1000	Snow field	−16½
6427	2000	Farm land and snow	−12
6448	1500	Farm land and snow	−11½
6474	1000	Farm land and snow	−12
6540	2000	Conifers and snow	−12½
6570	1500	Conifers and snow	−12½
6600	1000	Conifers and snow	−10
6714	2000	Conifers and snow	−9
6757	1500	Conifers and snow	−9
6778	2000	Smooth water	+4½
6871	3000	Snow	−5½
6927	2000	Snow	−4½
7063	3000	Snow	−7
7103	2000	Snow	−11
7165	1500	Snow	−6½
7216	1000	Snow	−8½
7263	3000	Conifers and snow	−13½
7297	2000	Conifers and snow	−11½
7329	1500	Conifers and snow	−12

7361	1000	Conifers and snow	−14
7590	3000	Conifers and snow	−13½
7616	2000	Conifers and snow	−14
7691	1500	Conifers and snow	−15
7788	1000	Conifers and snow	−14½
9074	1000	Conifers, dry	−14
9148	2000	Conifers, dry	−12
9407	3000	Conifers, dry	−14
9763	1000	Sea, state 1	+½
9812	2000	Sea, state 1	0
9926	1000	Marsh	−12
9985	2000	Marsh	−12
10106	3000	Marsh	−10½
10128	1000	Coastal dunes	−8½
10549	1000	Farm land	−10 to −15
10704	2000	Farm land	−7½
10765	1000	Town (Worcester)	−6
10964	2000	Exmoor	−7½
10985	1000	Exmoor	−8
11194	2000	Sea	+2
11214	1000	Sea	+4
11402	1000	Sea-shore sand	−5
11964	2000	Sea-shore sand, low portion	−10
12070	1000	Sea, state 3 to 4	+2½
12087	2000	Sea, state 3 to 4	+2½
12147	3000	Deciduous forest	−16
12175	2000	Deciduous forest	−14
12233	1000	Deciduous forest	−15
12293	1000	Sea, state $<\frac{1}{2}$	+2½
12309	2000	Sea, state $<\frac{1}{2}$	+3½
12460	1000	Flat sand	−8½
12474	1000	Over edge of a dune	−6
12486	1000	Flat sand with pimples	−3½
12499	1000	Flat sand with pimples	−5
12512	1000	Flat sand with pimples	−4
12520	1000	Across confused dunes	−7½
12521	1000	Across confused dunes	−26
12540	1000	Over confused dunes	−23½
12567	1000	Flat sand	−5½
12594	1000	Confused dunes	−13½
12605	1000	Flat sand	−6
12635	3000	Flat sand	−6
12676	3000	Flat sand with pimples	−11½
12689	3000	Top of a dune	−23
12710	3000	Flat sand with pimples	−3½
12727	3000	Confused dunes	−4
12752	3000	Small confused dunes	−7
13355	1000	Curving dune	−31
13577	3000	Curving dune	−27½
13622	3000	Flat sand	−12½
13653	1000	Flat sand	−3½
13671	3000	Sea	+3½
13684	1000	Sea	+1½
13781	3000	Mostly dunes	−17
13827	3000	Mostly dunes	−16
13894	3000	Mostly dunes	−12

TABLE 2—*continued*

13904	3000	Mostly dunes	−15½
13910	3000	Mostly dunes	−19
13954	3000	Mostly dunes	−16
14054	3000	Mostly dunes	−15
14071	3000	Mostly dunes	−15½
14089	3000	Mostly dunes	−17½
14096	3000	Mostly dunes	−11
14132	3000	Mostly dunes	−11
14190	3000	Semi-desert S. of El Adem	−5½
14252	4000	Semi-desert S. of El Adem	−5½
14295	1000	Grey rock and large dune	−24
14323	1000	Grey rock and large dune	−25
14365	3000	Grey rock and large dune	−18
14406	3000	Grey rock and large dune	−17
14441	3000	Flat sand, N. of rock, (Specular)	−1[a]
14442	3000	Flat Sand (0.2 sec. duration)	−14[a]
21169	1500	Valley	−14½
21301	1250	Hillside	−17½
21543	2000	Town (Wolverhampton)	−5
21585	1000	Town (Wolverhampton)	−5
22294	1000	Valley, continuous 'A' scope	[b]
22338	1000	Valley, continuous 'A' scope	[b]
22341	1000	Valley, continuous 'A' scope	[b]
22607	2000	Sea, state 2, continuous 'A' scope, aircraft banked	[b]
23105	2000	Deciduous forest	−16½
23145	1000	Deciduous forest	−14½
23167	1000	Sea, state 2 to 3	+½
23182	2000	Sea, state 2 to 3	+1½
23861	1000	Factory	−7
23881	1000	Factory	−4
23898	1000	Factory	−8½
23912	1000	Factory	−4

[a] These readings of −1 and −14 db occurred adjacent to one another, −14 db was maintained for 0.2 seconds only.

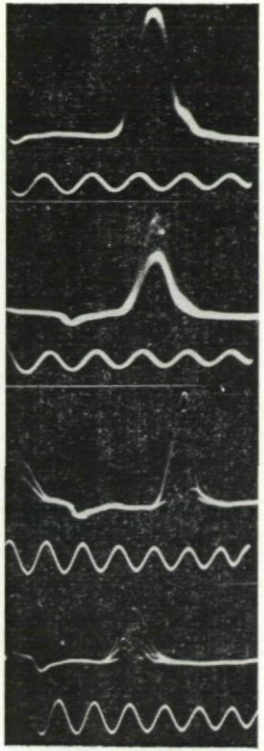
[b] Photographs are reproduced in Figs. 5 and 6.

It had been expected that low reflection gains would be observed in really cold weather when moisture in the lower parts of trees and to some depth in the ground would be completely frozen. However, no significant change was noted at the lowest temperature at which measurements were made (− 8 °C); presumably a much lower temperature is required to give these conditions.

8. *Snow Fields*

The reflection from this type of terrain is primarily by scattering, but the large range of block averages shows that considerable local variations in the terrain or snow conditions can occur.

It is theoretically possible for there to be very low reflection from a snow-covered surface as freshly fallen snow may have a density less than 0.1 and if the snow lies in layers gradually more and more compacted, a progressive change of density can occur with progressive absorption of the incident



3a. *Very smooth water* (6131).

3b. *Water, slight swell* (6239).

3c. *Rough water* (13684).

3d. *Forest* (9074).

FIG. 3. "*A*" *Scope Photographs. Each photograph is the superposition of* 10 or 20 *echoes. The timing wave is at* 4 *Mc/s.*

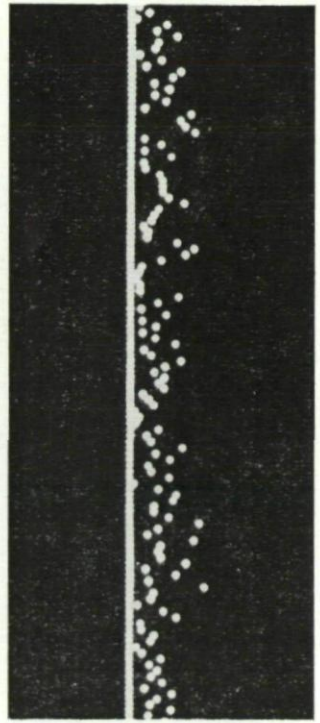


FIG. 4. *Normal method of recording echo amplitudes* (9074)

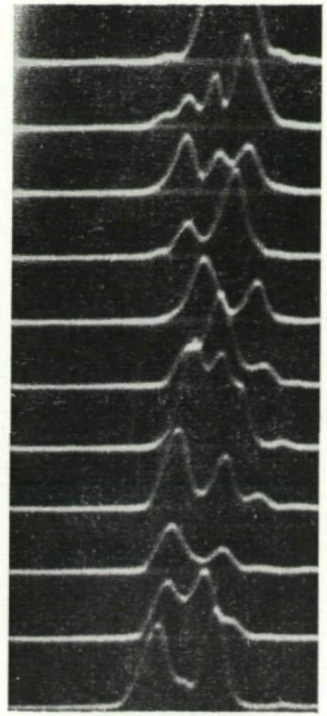


FIG. 5. *Steep valley* (22341). *Continuous 'A' scope record. Time base length* 1.8 *μs* P.R.F. 250/second

signal, which will reduce the reflection to perhaps −20 db. No such phenomenon was observed, and the mean reflection gain recorded was −9.5 db. The reason for this is thought to be that the snow was crusted on some occasions when measurements were made, and not deep enough on others.

9. *Urban*

Few measurements were made in built up areas, but it was noted that the instantaneous echo could vary very rapidly, for example over a factory roof the amplitude changed by 10 db in 1/10th second, although the mean echo over a block of 500 pulses was not very different from the typical value.

10. *Desert Sand*

Using the known composition of the sand in the Libyan desert over which measurements were made the reflection gain can be calculated to lie between −10 and −25 db assuming that reflection is by scattering. In fact the measured mean value was −5 db, and it was found that the reflection was primarily specular from the smooth sand.

Considerable local variations in mean reflection occurred: this is probably common in cases where reflection is mostly specular.

11. *Sand Dunes*

The type of dune over which measurements were made is the "Sief" dune. Despite the fact that the area chosen was one of the driest in the world occasional groups of "terabils"* were observed. As the results for flat sand show, the surface is sufficiently smooth to give specular reflection, so that for sloping sand wide variation in the mean reflection occurred; the extremes of the block averages were −4 and −31 db. The variations were often rapid.

If, as assumed, the extremely low amplitude returns were due to the fact that specular reflection is directing most of the reflected energy away from the receiving aerial, the return should consist principally of residual scattered energy, and this was found to be the case.

12 *Non-Flat Terrains*

The terrain types listed in Table 1 were mostly selected for their homogeneity and flatness, but the effects of rough and undulating country are important in practice. Non-flat terrain can be divided into two classes:—

(a) Where the surface is a specular reflector. Examples of this type of surface are provided by sand dunes, discussed above. The characteristic of this type of surface is a return varying widely in amplitude.

(b) Where the surface reflects by scattering. This case was investigated in three ways; by simply banking the aircraft when flying over the sea, by flying along the sides of steep hillsides, and an extreme case by flying along the line of a cliff. It was found that the effect of the

* A terabil is a conical mound of earth a few feet high held in place by the roots of some desert plant such as the tamarisk.

sloping surface was to modify the echo shape and duration. An example of this is shown in Fig. 5 which shows a succession of received echoes from a steep hillside. It will be seen that the echo appears to come from a variety of ranges, multiple peaked echoes being common. The reason for this is because although when the terrain is level the spread of ranges within the −3 db points of the beam is only 40 feet, when the terrain slopes at 20° the spread is 300 feet, leading to the spread in range of echoes.

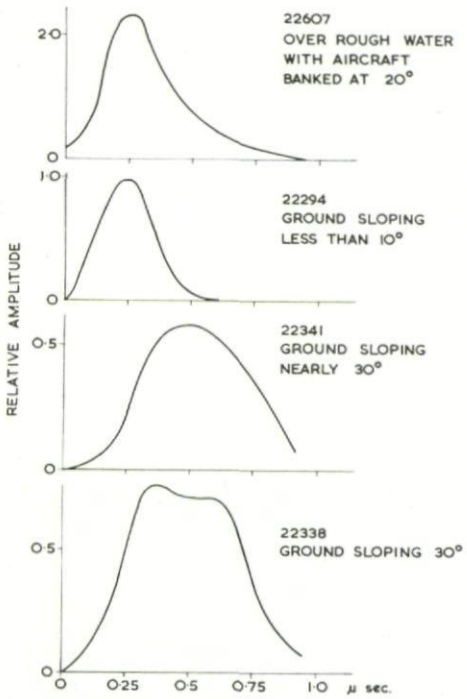


FIG. 6. *Average echo pulse shapes.* (*The zeros of the time scales are arbitrary*).

An "average pulse-shape" has been synthesised by dividing each echo into small time intervals, averaging the amplitude in the same interval of every pulse over a period and plotting the composite pulse shape obtained. Four such average pulse shapes are shown in Fig. 6. The measurements along the cliff showed, as would be expected, two distinct echoes.

The accurate calculation of the extension of the echo pulse from a sloping scattering surface is complex. To obtain a simple method of predicting approximately the expected extension of the echo, the measured pulse shapes were compared with the extension obtained simply by considering the spread of ranges within the −3 db points of the beam. The maximum pulse extension was found to be approximately equal to this calculated extension and the average extension about 75% of it.

7. CONCLUSIONS

The character of the echo from the sea and ground has been measured. The mean value, and spread, for each terrain type is shown in Table 1. Most surfaces were found to reflect by scattering, but some, very smooth water and dry sand gave specular reflections. The distribution of echo amplitudes from scattering surfaces has been found, as would be expected, to be closely Rayleigh. The reflection from inclined surfaces depends on whether the surface reflects specularily or by scattering. If specularily, very low amplitude returns occur and the local variations in amplitude is great. If by scattering, the effect is primarily to extend the duration of the returned pulse. The degree of extension has been found to be that expected from consideration of the aerial polar diagram.

DISCUSSION

A. S. WESLEY: Your work was all carried out on the S-Band. Can you give us any guide as to how your figures can be extrapolated to other frepuency bands?

A. R. DOMVILLE: I imagine that at shorter wavelengths more surfaces would tend to be scattering surfaces than is the case at S-Band.

CHAPTER 8

SOME INFORMATION ON NATURE, SIZE AND VELOCITY DISTRIBUTION OF RANDOM SCATTERERS GATHERED BY A PULSE RADAR ON TWO WAVELENGTHS USING VERTICAL FIXED BEAMS

W. Fogy
Institut für Flugfunk und Mikrowellen, Oberpfaffenhofen, Germany

In recent years the D V L developed a vertical-beaming cloud-radar, operating on two wavelengths simultaneously (X- and K_a-band). The two postdetection outputs, after sampling with range gates, are digitalized and stored for several seconds on a fast running film at a rate of 7000 bits per second per channel. The information contained on the film is transformed into punched tapes and fed into a computer for calculating the two first-order-distribution-density functions, two autocorrelation functions and one cross-correlation between the two channels. On the basis of these five statistical descriptive functions an attempt is made to read out information about the nature and some statistical parameters of clouds, rain, snow by comparison with other meteorological data, especially those measured at nearly the same time by means of an aeroplane within the range-gated pulse volume under test. Some results of a series of experimental runs are presented and discussed.

1. INTRODUCTION

Normal pulse radar became in the last two decades a powerful tool for the detection, visual presentation and quantitative measurement of atmospheric phenomena such as precipitations, clouds, moisture convections (Angels), and clear air turbulence (CAT). One of the types widely used in this field is the vertical-beamed radar with its ability to investigate the height-time echo structures, especially of clouds and "Angels." Since the optimum wavelength cannot be determined uniquely without considering the destination of the radar and its system application, the basic idea of our work was to look at the same hydrometeorological event simultaneously with two radar wavelengths, in order to compare the results with each other and if possible with flight measurements. Measurements of natural events bear the basic disadvantage that an individual measurement is never repeatable under identical conditions. Therefore the quantity of material must be high, and it is most important to have stable equipment which can be calibrated. To develop this is of course a long term programme by itself and

a high percentage of the effort had to be directed towards getting the equipment into stable and precise operation.

1.1.

Cloud parameters are: Drop size distribution, number of particles per unit volume nature of particles (ice, water or mixed phase), and the velocity distribution of these particles. The measurement of these parameters and their variability with space (height) and time by the use of radars is a very important study and numerous scientists have made important contributions during the last 20 years [Siegert, Martin, Wallace, Atlas, Austin, Fleischer, Hilst, Hitchfeld, Rogers, Lhermitte et al.]. In a non-coherent pulse radar with pulse length θ, the instantaneous echo signal gated at a fixed range (delay T) represents the vector sum of all individual particle returns, coming from the range difference

$$c\frac{T-\theta}{2} \quad \text{to} \quad c\frac{T}{2}$$

weighted with the gain function of the radar beam. The phases of the partial vectors will change from pulse to pulse, corresponding to the radial movement of the particles, the amplitudes will change due to the appearance, disappearance, growing, evaporating and splitting of particles within the pulse volume under test. If particles are present with non-spherical shape, aspect-changes may cause additional variations of amplitudes and phases.

As a result, both amplitude and phase of the sum vector will in general change simultaneously from pulse to pulse. Individual amplitudes and relative radial velocities are responsible for the resultant amplitude fluctuations (envelope), also in addition to the absolute radial velocities they are responsible for the resultant phase fluctuations (Doppler shift and spread[3,7,14]). These fluctuations are best described by their amplitude distribution density (DD) and their autocorrelation function (AKF).

1.2.

The dependence of both functions upon radar wavelength is of interest. Movements of particles are measured in terms of wavelength, therefore that part of the fluctuation which is caused by relative movements alone, must show in the AKF a decay time proportional to the wavelength used. If movement is not the only reason for fluctuations, deviations from proportionality may be expected and can be used as a source of additional information. On the other hand, if the radar cross section of all scatterers can be presented as a product of two functions, the one of them containing the wavelength alone, as is the case for small spherical droplets in the pulse volume only (λ^{-4} law), the wavelength dependence may be taken over in the radar equation to the "radar performance" and no additional information results from the distribution measurement. In general however especially looking at clouds, fluctuations are caused by movements as well as by cross-sectional-wavelength dependence, so that both "DD" and "AKF" deviation from the normal case might be expected.

1.3. *Distribution Density-function*

Investigations of the instantaneous power distribution showed under assumptions of random positions from pulse to pulse, sufficient short range gate, sufficient high S/N-ratio, sufficient large number of particles within the pulse volume under test and a sufficient large number of single observations that a Raleigh distribution with an average return power as the only descriptive parameter will in general lead to quite fair agreements with measurements[1,2,5,6]. The radar wavelength used is not reported to influence the type of distribution. However considerable deviations from Raleigh distribution have been observed with different equipment at different wavelengths, but have been interpreted as being due to equipment errors rather than an informative deviation from the Raleigh distribution. Most reported distribution measurements have been made against rain, snow, hail, but little is known relating to clouds.

A second important result in connection with amplitude distribution as a source of information was found in the so called Z-R (Z-W resp.) relation[5,15,16,17]. That relation gives a general linear two parameter regression between the logarithm of the radar reflectivity Z (a measure for the average return power or the first moment from the DD) and of the liquid water content per unit volume. (R in case of rain, W in case of clouds.) From particle size distribution, density and nature of scatterers the value Z can be calculated and compared with the radar measurement directly. From such measurement considerable variations of the two regression parameters are reported from different investigators using different equipments at different times and places. (A brief outline is presented by Scolnik.[2])

From these results neither the particle density, nor the size distribution, or nature of particles can be derived uniquely. The basic idea of our measurements is therefore, to try to get additional information from distribution measurements using two wavelengths simultaneously and investigating their deviations per wavelength using Raleigh distribution and Z-W relationships.

Let us assume that only spherical water droplets with a given size distribution fill the pulse volume homogenously, the maximum diameter being small compared with the shorter wavelength. We are then in the Raleigh-scattering region and the two Z-values are inverse proportional to λ^4. If the performance of equipment is known, the measured average return powers are specified. However considerable deviations from the 4th power law are reported and have been observed with our own equipment. The reason for these deviation can be as well a nonhomogeneous space distribution of radar reflectivity within the pulse volume, as the presence of a sufficient large number of scattered reflectivity deviating from the Raleigh law at least at one wavelength. A smaller pulse volume combined with higher efficiency of the radar might overcome such a difficulty.

Furthermore the amplitude distribution depends on relative motion of the scatterers toward or away from the radar. If we assume the unrealistic model with no such relative motion (all velocity vectors being the same), the

result would be a narrow-peaked amplitude distribution, because we have no fluctuations of the signal envelope, with small variance about the average power like a shifted gaussian. The later depends not only on number and size distribution of scatterers, but also on their space distribution. Such an amplitude distribution can be expected for relative short times and is indicative for a low degree of relative motion, but not indicative for the value Z, because the space distribution is not known.

We assume now another hypothetical model with a large number of equal small scatterers with a random relative velocity distribution giving from pulse a random position of each scatterer. Theory predicts the instantaneous return power Raleigh distributed with the first moment proportional to the sum of the individual return powers. Such a distribution can be found only for a relatively short time and is indicative as well for the value Z as for the presence of turbulence high enough to produce independence between consecutive pulses.

These two examples show the boundaries between which amplitude distributions are to be expected. In realistic cases with realistic drop-size distribution, the motion depends on the inertial characteristics of scatterers within the air velocity and the gravity field. Actual distributions can therefore be expected to be found anywhere between the two previous mentioned boundary distributions, the one a Raleigh distribution, the other a shifted Gaussian distribution. There exist two families of such transient distributions: the m distribution of Nagakami[23] and the Rice distribution [a sinusoidal carrier enbedded in narrow band noise], sometimes called excentric Raleigh distribution.[20,22] The first one is in general better matched to our problem, therefore we used the m-distribution to describe cloud echo amplitude DD functions.

1.4. *The Autocorrelation Function*

The second order statistics of the envelope fluctuations lead to the autocorrelation function of the return amplitude. Using noncoherent radar, Doppler shift and spread are not available as source of information, but pulse to pulse envelope time fluctuations can be utilised. These are caused directly by relative motions of individual scatterers perpendicular to the phase front of the beam. In section 3 a brief outline of known relations concerning fluctuation spectra and correlation functions is presented.[3,7] As far as all these theories make use of Wiener-Kinchin theories and give a spectral interpretation, they assume the observed process to be stationary, which is seldom the case with clouds, if observation times more than a few seconds are used. The dilemma is, that at one hand we require several thousand individual echo data from each radar, which can deliver at best a few thousand data per second. On the other hand an event does seldom hold stationary for longer than half a second. If however this can be assumed, the Fourier transform of the autocorrelation function of amplitude represents the relative velocity distribution density under the assumption that drop size distribution is known and is homogeneous over the pulse volume under test.[7,...,13] In case of clouds and vertical beaming radar however these assumptions hold very

seldom. Therefore the application of simplified theories is very disencouraging and the best we could do was to gather material for AKF's at two wavelengths to compare them with direct measurements.

2. THE *Z-W* RELATION

The *Z-W* relation has been studied by a number of experimenters with rain and clouds as a target. Starting from the radar equation for a single spherical scatterer as a point target with radar cross-section σ_i, we get the return power

$$P_i = \beta_i^2 \frac{P_t G_i^2 \lambda^2}{(4\pi)^3 R_i^4} \sigma_i \qquad \text{with} \tag{1}$$

$$\sigma_i = \pi^5 |K|^2 \frac{D_i^6}{\lambda^4} \qquad \text{for a dielectric sphere} \tag{2}$$

with diameter $D_i \ll \lambda$, $|K|^2 = 0.93$ for water (0.197 for ice) and β_i a one way attenuation factor between transmitter and target, caused by scattering and dissipation.

All echo powers will be summed over the pulse volume. Generally this is done by replacing the individual values β_i, G_i and R_i by average values $\bar{\beta}$, $\bar{G}$, and $\bar{R}$. A value $Z = \sum_{i=1}^{N} D_i^6$ is defined, taking the sum over all particles per unit volume. The pulse volume under test can be expressed in terms of gain, wavelength and pulse length τ. Using the signal-to-noise power ratio $(\bar{S}/N)$ at the receiver input and a reference distance R_0 in the form $\bar{R} = \frac{\bar{R}}{R_0} \cdot R_0 = \bar{\rho} R_0$, one gets

$$\left(\frac{\bar{S}}{N}\right) = \mathscr{P}_{\text{radar}} \cdot \left[\frac{\bar{\beta}^2 |K|^2}{\bar{\rho}^2}\right] Z \tag{3}$$

with $\mathscr{P}_{\text{radar}}$ a constant describing the radar performance

$$\mathscr{P}_{\text{radar}} = \left(\frac{\pi}{4}\right)^4 \frac{P_t \bar{G} c \tau}{\lambda^2 P_{\text{noise}} R_0^2} \tag{4}$$

The reader interested in details, is referred to the literature.[1,2,5,17] Between Z, the "radar reflectivity" per unit volume, and the liquid water content $W = \frac{\pi\gamma}{6} \Sigma D_i^3$ have been developed[5] the following relationships

$$Z = \sum_{i=1}^{N} D_i^6 = N \sum_{K=1}^{D\max/\Delta D} \nu_K D^6 \Rightarrow N \int_0^{D\max} D^6 p(D)\, dD \tag{5}$$

$$W = \sum_{i=1}^{N} D_i^3 \cdot \frac{\pi\gamma}{6} = \frac{\pi\gamma}{6} N \sum_{K=1}^{D\max/\Delta D} \nu_K D_K^3 \Rightarrow \frac{\pi\gamma}{6} N \int_0^{D\max} D^3 p(D)\, dD \tag{6}$$

$$\frac{Z}{W} = [D_{\max}^3] \cdot \left(\frac{6}{\pi\gamma} \frac{\int_0^1 x^6 p(x)\, dx}{\int_0^1 x^3 p(x)\, dx} \right) \tag{7}$$

or in another form

$$\frac{Z}{W} = [D_0{}^3] \cdot \left\{ \frac{6}{\pi\gamma} \cdot \left(\frac{D_{\max}}{D_0}\right)^3 \frac{\int_0^1 x^6 p(x)\, dx}{\int_0^1 x^3 p(x)\, dx} \right\} \tag{8}$$

$$\gamma = \text{spec. weight}$$

$$x = \frac{D}{D_{\max}} = \text{rel. Diam.}$$

$$p(x)\, dx = p(D) dD$$

$$\int_0^1 p(x)\, dx = 1$$

D_0 is some averaged diameter. Following Atlas, it was proposed to use $D_{W50\%}$ as a most descriptive and representative median-diameter. That is the diameter up to which 50% of the liquid water content W_{50} is reached. Many observed materials (rain and clouds) have been investigated, calculating Z and W for each droplet spectrum. A regression of the form:

$$A = aW^b \tag{9}$$

has been found, in which for the different phenomena (rain and clouds) the parameters a and b show large variability, especially for different cloud types. The great variability is due to the statistical behaviour of the two factors in Eqs. (7) and (8), and their statistical degree of dependence. Errors can be very large in measuring Z by radar and deducing a value for W, even if the radar is very stable and precise, because the actual droplet size distribution, the degree of homogeneity inside the pulse volume and the dynamic behaviour of the droplets under air dynamics are not known. It is clear that the measurement of average power alone is not sufficient to determine Z more accurately.

It has been important, to calculate corresponding values for Z and W from measured natural droplet size distributions[5] and to investigate their statistical behaviour over a variety of different clouds.

Only the log of Z and W are of interest in these experiments. We write therefore the Eqs. (7) and (8) so as to give the relationship between log Z and log W

$$10 \log Z = 10 \log W + \underbrace{10 \log [D_{W50}^3]}_{\sigma_\mu 50} + \underbrace{10 \log \left[\frac{6}{\pi\gamma} G_{50}\right]}_{\Gamma 50} \tag{10}$$

or the second form

$$10 \log Z = 10 \log W + \underbrace{10 \log [D_{\max}^3]}_{\sigma_\mu} + \underbrace{10 \log \left[\frac{6}{\pi\gamma} F\right]}_{\varphi} \tag{11}$$

$$\text{with } G_{50} = \frac{6}{\pi\gamma} \cdot \frac{1}{x_{50}{}^3} \cdot F \tag{12}$$

being a form factor of a given droplet size distribution supposed by Atlas and

$$F = \frac{\int_0^1 x^6 p(x)\, dx}{\int_0^1 x^3 p(x)\, dx} \tag{13}$$

being a relative form factor, defined as the ratio of the 6th moment to the 3rd moment of the normalized droplet size density distribution function.

$$(x_{50} = D_{W50}/D_{\max}).$$

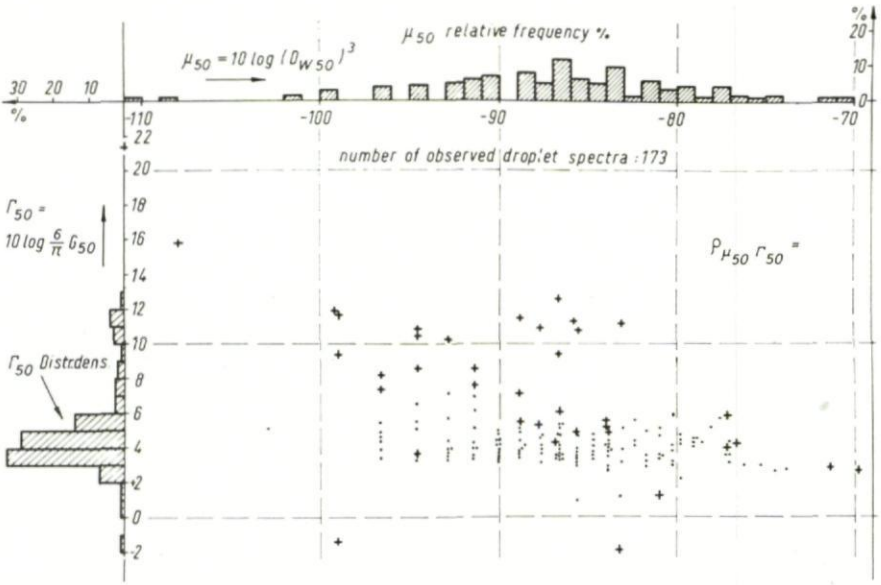


FIG. 1. *Relative frequency of* μ_{50} Γ_{50} *and regression for observations in flight* (134 *Diem observations marked* ·, 39 *by author marked* +)

$D_{\max}$ depends on the method of measurement, but can be thought of with good approximation as D_{W99}.

Figures 1 and 2 show distribution histograms of μ_{50}, Γ_{50} as defined in Eq. (10) ($\mu_1\varphi$ from Eq. (11) resp.) as well as their statistical dependence. The values have been calculated from 39 of our own and 134 "Diem"-flight observations on droplet size distributions. The figures show very clearly the statistical independence of μ and φ (the correlation coefficient $\rho_{\mu\varphi}$ between μ and φ being less than 0.05) and a regression between μ_{50} and Γ_{50}, as expected. Furthermore a comparison shows, that for use in radar measurements the decomposition of the Z-W relation into μ and φ by normalization with maximum diameter leads to a smaller variance for

both factors as compared with μ_{50} and Γ_{50}. On this point looking at Figs. 1 and 2 one notes, that surprisingly the spread of μ is less than that of μ_{50}.

The relatively small number of large droplets in a given spectrum contributes most to the radar reflectivity Z. Therefore the dynamic behaviour of these droplets will often control the fluctuations of the return power and is responsible for the high degree of regression uncertainty in Z-W cloud relations. This has been shown on our own observation material by eliminating the influence of $D_{\max}$ in the Z-W regression. The exponent b in the regression Eq (9) becomes unity and the correlation between the normalized values Z and W approaches unity (Fig. 3).

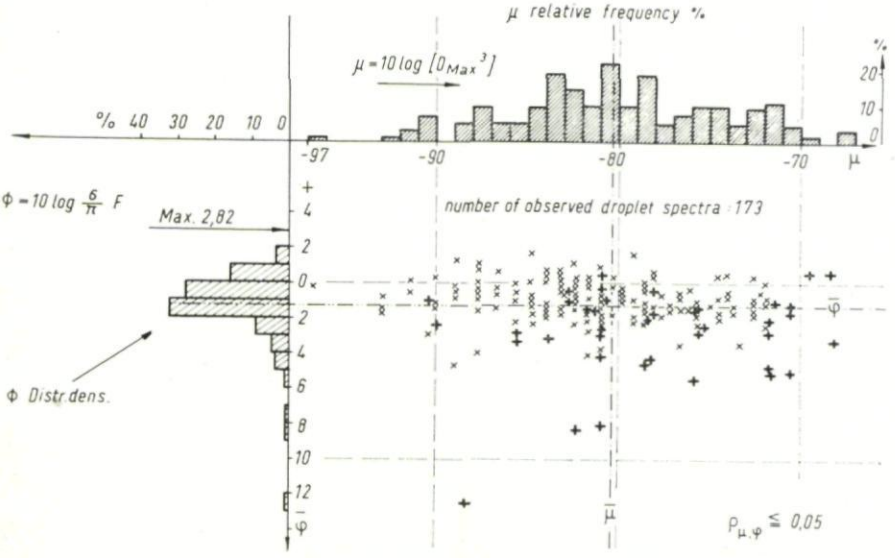


FIG. 2. *Relative frequency of μ, ϕ and regression for observations in flight* (134 *Diem observations marked* ×, 39 *by author marked* +)

It seems reasonable to look for additional information about droplet size distribution, especially for the upper part of it. On the basis of a non-coherent radar only, the measurement of the whole echo power distribution together with its autocorrelation function can contain such additional informations. (Measurements of Doppler spectrum, which have been made in recent years,[14] are not considered here.)

It is known, that cloud-events in a given pulse volume in most cases do not hold constant for a time long enough to gather sufficient sample-materials for a Z-statistic. Particle size distribution and/or velocity distribution changes with time. Records have shown, that often an observation time of several seconds is too long. This makes it difficult to measure only stationary

processes. The best we can do is, to take measurement-runs over observation-times, long enough to gather sufficient samples, but short enough to be free from marked change.

We expected, that if measurements of the amplitude distributions are made simultaneously on two radar wavelengths (3 cm + 8 mm), from these two distribution-functions we could gather more information about the particle content of the pulse volume under test, for instance D_{max}, and to use this additional knowledge for a better measurement of Z. It could be expected, that without simultaneous measurement of the corresponding

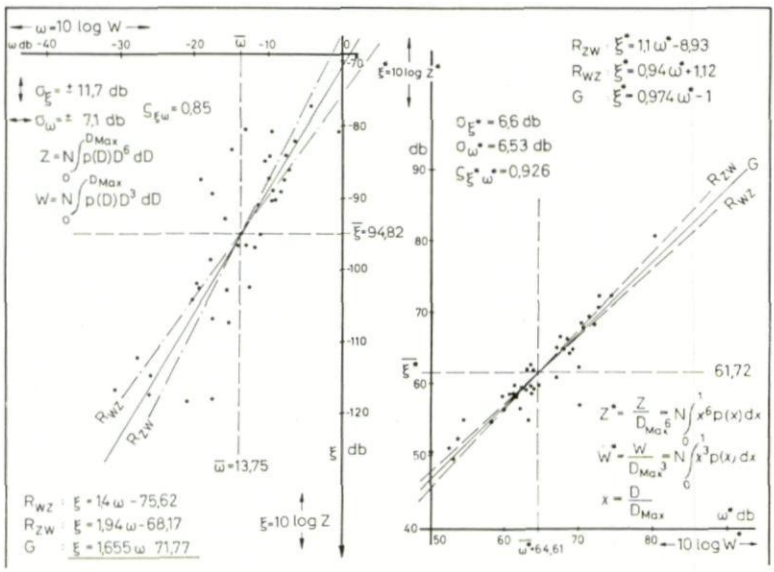


FIG. 3. *Regression Z versus W from* 39 *of authors observations, elimination of* D_{max}.

AKF on the two wavelengths a successful interpretation would not be possible, therefore our equipment was made to give both distributions and AKFs. In the field of statistical interference phenomena the generalized m-(Nagakami) distributions[23] may be descriptive for our problem, because this class uses two independent parameters, to describe transient distributions from the "one side" gaussian ($m = \frac{1}{2}$) through the classical Raleigh ($m = 1$) up to shifted "gaussian like"-distribution ($m > 2$). Figure 4 shows that class, in a presentation[24] deviating from that of Nagakami due to normalization with $\bar{x}$ instead σ_x.

Calibration measurements with noise as well as investigation of cloud signals seem to fit quite fair with m-distributions, more material will be needed before making such conclusions firmly.

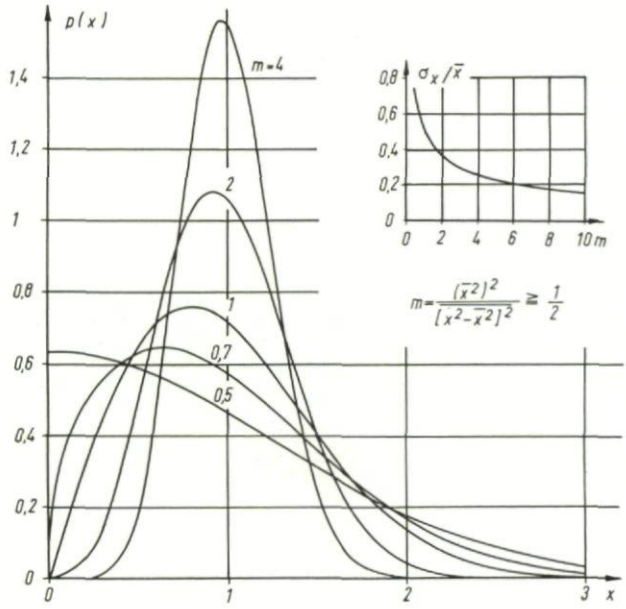


FIG. 4. *Generalized Raleigh (Nagakami) m-distribution normalized with the average value.*

3. THE AUTOCORRELATION FUNCTION (AKF)

Fleisher has given[7] a rigorous treatment of the information contained in the second order statistics. The envelope of the returned signal is expressed as the vector sum of the individual contributions

$$A(t) = \sum_{K=1}^{N} a_K e^{j\gamma D_K(t)} \qquad \gamma = \frac{4\pi}{\lambda} \tag{14}$$

with only D_K the distance of the kth scatterer from the radar, as a function of time. The effect of range inside the pulse volume has been ignored. The normalized AKF is then defined as the time average

$$G(\tau) = g(\tau)e^{j\psi(\tau)} = \frac{\overline{A^*(t) \cdot A(t+\tau)} - \overline{A(t) \cdot A^*(t)}}{\overline{A^*(t) \cdot A(t)} - \overline{A(t)} \cdot \overline{A^*(t)}} \tag{15}$$

$g(\tau)$ is even, $\psi(\tau)$ odd, from Eqs. (14) and (15) one has:

$$g(\tau)e^{j\psi(\tau)} = \frac{\overline{\left[\sum_1^N a_K e^{j\gamma D_K(t)}\right] \cdot \left[\sum_1^N a_e e^{j\gamma D_e(t+\tau)}\right]}}{\left[\sum_1^N a_K e^{-j\gamma D_K(t)}\right] \cdot \left[\sum_1^N a_e e^{j\gamma D_e(t)}\right]} \tag{16}$$

assuming uniform distributed phases of individual echoes, integral wavelengths in the range gate, and sufficient high number of particles, $\overline{A_K^*} = \overline{A_K} = 0$. Because the phases of different scatterers are independent at anytime the mixed product partial sums of Eq. (16) vanish and we have the AKF

$$g(\tau)e^{j\psi(\tau)} = \frac{\overline{\left[\sum_1^N a_K{}^2 e^{j\gamma v_K \tau}\right]}}{\left[\sum_1^N a_K{}^2\right]}$$

with

$$v_K = \frac{D_K(t-\tau) - D_K(t)}{\tau} \tag{17}$$

The radial velocity of the kth scatterer. Eq. (17) can be rewritten

$$g(\tau)e^{j\psi(\tau)} = \frac{\frac{1}{N}\sum_1^N \overline{a_K{}^2 e^{j\gamma v_K \tau}}^t}{\overline{a_K{}^2}} = \frac{\overline{\overline{a^2 e^{j\gamma v\tau}}^t}^K}{\overline{a^2}} \tag{18}$$

the average over scatterers and time, which can be written in the limit

$$g(\tau)e^{j\psi t} \Rightarrow \frac{1}{\overline{a^2}} \int_0^\infty \int_{-\infty}^\infty a^2 p(a, v) e^{j\gamma v\tau}\, da \cdot dv \tag{19}$$

$p(a, v)$ being the joint probability density of individual amplitude and velocity, independent of absolute time t.

The AKF of the intensity $A \cdot A^*$ is shown to be well approximated by $g^2(\tau)$, the square of the module of $G(\tau)$. $\overline{a^2}$ represents the average signal power. $g(\tau)$ is the only information delivered by noncoherent radar. $\psi(\tau)$ can only be measured by using pulse-Doppler radar. Equation (19) indicates, that information about velocity distribution can be resolved, by using inverse transform methods if the two variables a, v can be separated. Separation can be attained if either independence of a and v, or analytical relation between them can be assumed. A theory for the case of statistical relation between a and v is not known to me. It seems useful to rewrite the scatterer radial velocity as the sum of its ensemble-time mean value $\bar{v}$ and the individual deviation from the mean, v', which is a statistical variate.

$$v = \bar{v} + v', \quad \bar{v} = \text{const.} \quad \bar{v}' = 0 \quad p(v)\, dv = p(v')\, dv'$$

$$p(a, v)\, da\, dv = p(a, v')\, da\, dv' \tag{20}$$

then Eq. (19) gives

$$g(\tau)e^{j[\psi(\tau) - \gamma\bar{v}\tau]} = \frac{1}{\overline{a^2}} \int_0^{+\infty} \int_{-\infty}^\infty a^2 p\, (a_1 v') e^{j\gamma v'\tau}\, da\, dv' \tag{21}$$

In general the mean as well as the relative velocities enter into the phase function $\psi(\tau)$, because the right side need not be real.

In the case of independence between (a, v'), Eq. (21) can be integrated in a with $p(a, v') = p(a) \cdot p(v')$, we get:

$$g(\tau)e^{j[\psi(\tau)-\gamma\bar{v}\tau]} = \int_{-\infty}^{+\infty} p(v')e^{j\gamma v'\tau}\,dv' \tag{22}$$

the right side is now the characteristic function of the relative velocity distribution density $p(v')\,dv'$.

If the left side is known, as we can expect with coherent radar, $p(v')$ can be resolved by using the inverse transform method[7] about $\gamma\tau/2\pi$.

$$p(v') = \frac{\gamma}{2\pi}\int_{-\infty}^{+\infty} g(\tau)e^{j[\psi(\tau)-\gamma\bar{v}\tau]}e^{-\gamma v'\tau}\,d\tau \tag{23}$$

which should be real and independent of radar wavelength, therefore the phase function $\psi(\tau) - \gamma\bar{v}\tau$ should be even too. If only $g(\tau)$ is available, additional restrictions result. Only and if only $p(v')$ is an even function, the right side of Eq. (22) remains real. This leads to the same condition on $\psi(\tau)$ as does Eq. (23).

$$\chi(\tau) = \frac{4\pi}{\lambda}\bar{v}\tau \pm n\frac{\pi}{2} \tag{24}$$

The phase part $\psi(\tau)$ of the AKF represents then the mean Doppler-shift, relative velocities do not enter in it, and

$$\overset{\text{even}}{p(v')} = \frac{2\pi}{\gamma}\int_{-\infty}^{+\infty} g(\tau)e^{-j4\pi/\gamma v'\tau}\,dt \tag{25}$$

Since $g(\tau)$ by definition is even, the right side of Eq. (25) is real. In Eq. (23) the exponent of the phase function must be even too. For this case Fleisher shows, that the condition that $p(v^1)$ must be a probability density function, $(S\gamma(v')\,dv' \Rightarrow 1)$, leads to Eq. (24) as well. From these conditions follows: symmetry of $p(v')$ is the necessary and adequate condition for getting real AKF; which is inversely related to the variance of $p(v')$. If symmetry is not fulfilled, neither independence between a and v' should be assumed, nor should a non-coherent radar be sufficient for adequate information. Because with coherent radar $(\psi(\tau) - \gamma\bar{v}\tau)$ represents Doppler shift and spread, from the symmetry condition in Eq. (23) it follows that a coherent radar should measure symmetrical spread about the mean Doppler shift. Unfortunately Doppler radar was not available for these trials and only in recent years has Lhermitte[14] published the first results. Autocorrelation functions, measured simultaneously from the same pulse volume at different wavelengths, should agree between each other, if τ is normalized with λ. Differences may indicate that our assumptions do not hold (which is to be expected when observing clouds with radar).

The main assumption among all restrictions from the proceeding outline has been the independence between a and v'. This can be true in case of horizontal beamed radar, but not when vertical beams are used. For

instance if we look at only small water droplets in quiet air, the "terminal velocity" law is dominant and a could be uniquely defined by a function of radial velocity v. Then from Eq. (18) we get the AKF[7]

$$\frac{\overline{a^2(v)e^{j\gamma v\tau}}}{\overline{a^2}} = \int_{-\infty}^{\infty} \frac{1}{a^2}\, p(v)a^2(v)e^{j\gamma v\tau}\, dv \tag{26}$$

It depends now upon the function $a(v)$ whether the decomposition $v = \bar{v} + v'$ will be applicable or not.

In case $a = a(\bar{v})$ alone and $\bar{v}$ is constant over ensemble and time of observation, we can proceed as before and get

$$g(\tau)e^{j\psi(\tau)} = \frac{a^2(\bar{v})e^{j\gamma\bar{v}\tau}}{\overline{a^2}} \int_{-\infty}^{+\infty} p(v')e^{j\gamma v'\tau}\, dv' \tag{27}$$

again $\psi(\tau) - \gamma\bar{v}\tau$ must be zero or symmetrical as before but we must know $\bar{v}$ and $a/\bar{v}$, in order to determine the relative radial velocity distribution. With $a(v)$ we must know the nature and size distribution of scatterers. The mean radial velocity is by the Z-W relation correlated with $\overline{a^2}$ and within the frame of this one could resolve for $p(v')$. (The case of only functional relation to relative velocities $a = a(v')$ is very unprobable.)

In general however we have to allow for an unknown non-linear statistical regression between a and v because of the very complicated dependences between particle size and nature in their airstream, gravitational, temperature pressure and moisture-environment. No such theoretical treatment is known to me. The problem whether or not it is possible to split up the regression into an analytical and an independent statistical relation should be investigated. In general between a,v exists in addition a functional dependence of λ too, therefore the measurement of AKF on different wavelengths might deliver additional information, if the strong restrictions made before do not hold. Only experiments on known events can, so long as an adequate theoretical treatment is not known, clear up the situations. For instance Hitschfeld[7] presented some statistical material over the normalized decay time of AKF for different measurements on different wavelengths. These time constants show a very wide variability from 1 to 10 ms/cm wavelength. This cannot be explained fully.

1. *Crosscorrelation for two wavelengths (KKF)*

M. Vogel has shown[24] that the fluctuations of instantaneous radar return using two spaced carriers, become less and less dependent, if the frequency spacing is wide enough and the number of scatterers sufficiently high. The crosscorrelation at delay time $\tau = 0$ for instantaneous values is therefore to be expected to be very low whether the pulse volumes for both carriers are equal or not. Using relative short observation times and different but coaxial pulse volumes, as we did, the crosscorrelation should give some idea about homogeneity of space distributions of radar cross section per unit volume.

4. RADAR SYSTEM EQUIPMENT

A description of our cloud-radar equipment has been published.[16,18] The equipment has been modified and is now more powerful and stable in operation. The radar works with two spaced parabolic antennae, one for transmission, the other for reception, both operating on two wavelengths (3.2, and 0.86 cm). The parabolic dishes (diameter 1.75 m) are designed with the focus a little inside the aperture in order to achieve, together with an appropriate aperture illumination, a sufficiently low side lobe level. This resulted in good resolution of weak cloud echoes in the main lobe against aeroplanes or buildings on higher order side lobes and gives a very low leakage between transmitter and receiver. Technical data are:

	$\lambda_1 = 0.86$ cm	$\lambda_2 = 3.26$ cm
Antenna parabolic $\frac{f}{D} \cong 0.23$	equal 1.75 ϕ,	centre-distance 2.6 m.
Transmitter Pt kW	20—40	120—140
Pulselength τ μs, changeable		0.25, 0.5, 0.75
PRF c/s (λ_1 or λ_2)		1000 alternate
PRF ($\lambda_1 + \lambda_2$) c/s		2000
Gain vertical db	51	41
Beamwidth (halfpower)	1.8°	0.5°
IF		30 Mc/s common main amplifier
B_{IF}		$1 + 5$ Mc/s variable
B_{video}		$0 + 10$
Range gates (changeable)		0.5, 1.0 μs
Rectification		linear
Dynamic amplitude range		
before range gate	40 db	
after range gate	35 db	
Receiver Input: Matched Diode pairs		
Receiver noise power at input:	— 91.5 dbm	— 97.5 dbm
Radar Performance $\mathscr{P}_{radar}$[cm^{-3}] (Reference distance $P_0 = 1000$ m)	155 db	147 db

To avoid any nonlinearity, no AGC is used. The signals are matched to the dynamic range with IF-attenuators. For calibration purposes IF-white noise can be fed into each or both channels. A calibration in S/N ratio could be made quickly by visual comparison of the two periodically switched video noise traces on a A-scope, which is synchronized with the two radar pulses. This type of relative comparison is very accurate.

Our efforts to get stable operation on the 0.86 cm band have been very high. After elimination of sources of instability concerned with oscillators and silicon diodes, the remaining instability was in the Q-band magnetron. In recent years we have used French L T T-Magnetrons with medium

power, high pressure in waveguide system and have excellent stability in operation and life time, so that no AFC circuit is necessary and transmitter pulses show no significant AM.

The performance could be measured with an over all accuracy of ± 1 db. The antenna gain and diagram had been measured[16] before. The vertical direction of antenna axis presents a problem, which without a fixed target can be solved only by statistical methods over a long time period. There is

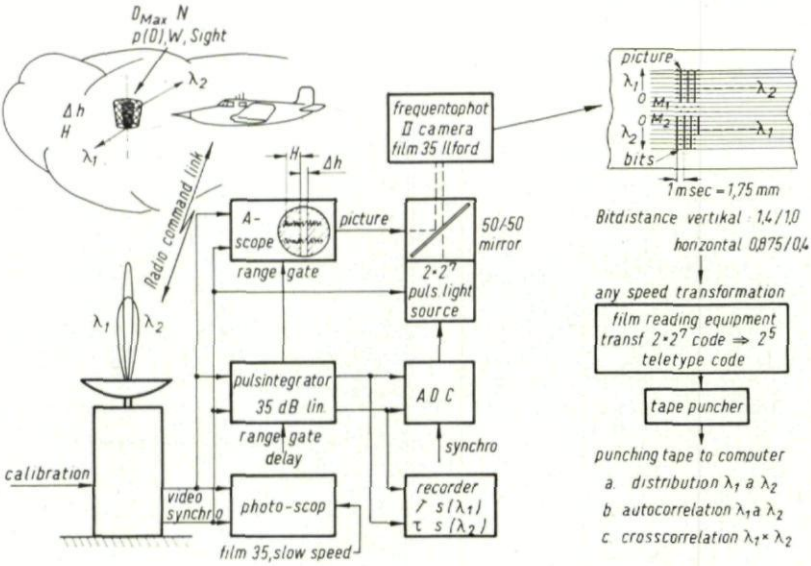


FIG. 5. *Data handling block circuitry for cloud radar*

some near field loss in effective antenna gain, but its influence is less than 3 db at heights more than 500 m and can be neglected over 800 m.

Figure 5 shows the data handling block circuitry we used in addition to the normal control A-scope. The following recording equipments are in use.

(1) The so-called *Photoscop*, widely in use with vertical radars. It gives on a very slow moving film on both channels, the height-time intensity structure of clouds or angels and represents by its long integration time on screen and film the most sensitive method of presentation. (S/N ratio less than unity can be detected.) The advantage of this type of registration lies in its sensitivity, low film consumption and ability to record for long periods without any human control. Its disadvantages, however, are its small dynamic range (less than 20 db), and the impossibility of using the cameras at a continuous slow film speed. Therefore an intensity modulation effect results, which makes quantitative automatic analysis unattractive.[16,17]

(2) *The radar pulse integrator (RPI), Fig. 6.*

It provides range gates for two gate lengths (0.5, 1 μs). The time position of the gate can be chosen either by hand in steps of gate length, or can be scanned linearly, step by step, with three different times (1, 10, 100 radar periods). The programming technique is digital. The gates operate on both channels and cut out from the video signal the corresponding piece and take the average value over the gate length, so that on each wavelength's output occurs a PRF-

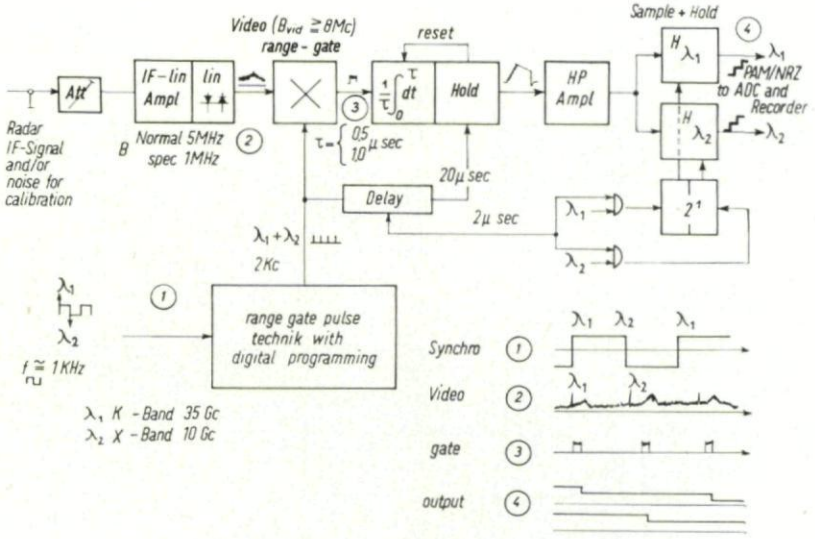


FIG. 6. *Circuitry of radar pulse integrator for two wave vertical cloud radar*

synchronized PAM-NRZ signal, which represents, for narrow band input signals, approximately the instantaneous envelope amplitude, and for wide band signals, the average envelope over the gate length. The Gate, integrator and main amplifier are again common for both channels in order to ensure low errors. Figure 7 shows the amplitude characteristic of the whole measuring channel (radar plus RPI).

As we wanted to get simultaneously distribution functions and AKF, it was found useful to digitalize directly the two outputs of the RPI before these were passed to the store.

The radar was located far away from our computer (700 km) and a real time use of computer was not needed at all. An electronic-optical film-reading equipment was developed, which transforms speed and provides on its output a code suitable for a normal teletype machine to put the data onto a punched tape.

Using film storage there was a possibility of storing between the bit lines the corresponding video signal. Analog film records of video weather signals from our radar showed the presence of very fast fluctuations within

the gate length though the S/N ratio was high. This indicates that sometimes signals are to be thought of as wide band noise signals, and all restrictions for the theoretical treatment no longer hold. Normally for a narrow band signal fed through our wide band channel one would expect no

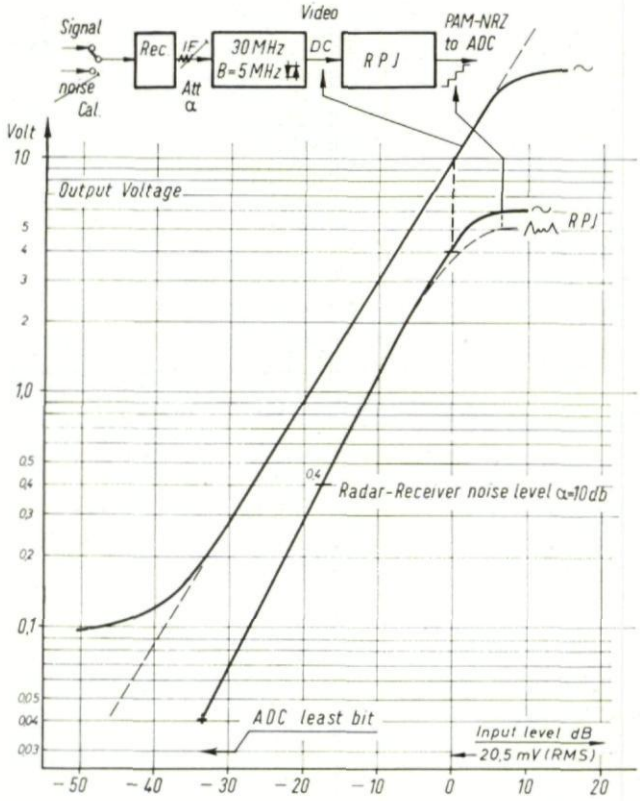


FIG. 7. *Amplitude characteristics of measuring channel*

fluctuations during the gate length. On the RPI output only fluctuations up to half the PRF can be defined uniquely, corresponding to the sampling theorem, and a pulse radar cannot smooth out higher informative spectral terms which are present in the target behaviour. Only a higher PRF or the use of CW-radar can help here. (For a cloud radar a PRF of 5 kc/s per wavelength was optimum but this would have been too expensive.)

5. SOME PRELIMINARY RESULTS FROM FLIGHT MEASUREMENTS AND CALIBRATION

A series of runs was taken in Nov. 1963. Unfortunately the films could not be utilized for our purposes. On the one hand the weather, was not right for our experiments (high gustiness and winds up to 70 km/h), and on the

other the high degree of echo dynamics and inhomogeneity of clouds made it impossible to make not only airborne measurements but also ground recordings. Also most of the films were tainted during automatic processing, so that only a few could be read. In April 1964 it was possible to take about 100 runs, some with and others without directly comparative flight measurements. The data processing of the films on our computer is still going on and no comparative flight data handling could be started. I can therefore present only a few results from the first processed films.

5.1.

First it was important to know, how calibration noise is processed through our channels since we use IF-band-limited white noise for calibration purposes. Noise runs were taken and AKF and distributions calculated. The AKF was found to decay within 1 ms for each channel. Several output distributions which we expected to be of Raleigh type for narrow band noise, ($m = 1$) following[6] are shown in Fig. 8, which compares the effect of wide band ($B = 15$ Mc/s at 30 Mc/s IF) and narrow (or middle) band noise (B = 1 Mc/s) on the RPI-output. For both cases runs were made with 0.5 and 1 μs range gate in order to show the influence of video averaging. We used two different noise levels, 15 and 0 db, the first using all steps (bits of ADC), the later only the first few steps. The Nagakami parameter m averaged over pieces of 256 values per channel [short time average] as function of time $\bar{m}^T(t)$ is shown in Fig. 9. The number N of values was in all cases several thousand, and proved to be sufficiently high. RPI-output short time averages $\bar{x}, y, \bar{\sigma}x, \bar{\sigma}v$ and $\bar{m}^T$ are calculated and their average values indicated. The results can be summarized:

(1) There is [Fig. 8] no Raleigh distribution for wide band noise. The measured histograms fit m-distributions. The Nagakami m-value for all wide band distributions is high (3–6).

(2) For narrow band noise m tends to unity, but the deviations from a Raleigh distribution are yet significant, so that it might be thought of as a "middle band noise" (m between 1 and 1.5).

(3) The averaging effect of range gate is seen clearly in case of narrow band noise. The m-value, however, is still 18% too high in the case of the short gate.

(4) There is a tendency for m to decrease with decreasing level in all cases, as seen on the right side of Fig. 8. The distortion of distribution goes down to $m = 0.8$. This effect was to be expected and can be used as an additional explanation for the deviations from Raleigh distributions in Austin's[6] work. The histograms shown there are sometimes only 16 levels used. This effect can be called "bit error."

(5) The short time average of $\bar{m}^T(t)$ [$T = 256$ ms and $\Delta t = 64$ ms] is presented in Fig. 9. The time fluctuations of the short time average of m for calibration noise are to be compared later with the corresponding value for a weather signal (Figs 9–12).

(6) Saturation effects do not seriously affect the result, if the degree of average channel saturation is below 4%. It is important to lock the signal before the run in such a way into the dynamic range of AVC that neither bit error nor saturation effects occur.

It is clear that additional noise material should be gathered, to get curves for the bit error as a function of "level in range" and the influence of bandwidth for different spectral shaped inputs on the DD-function. The influence

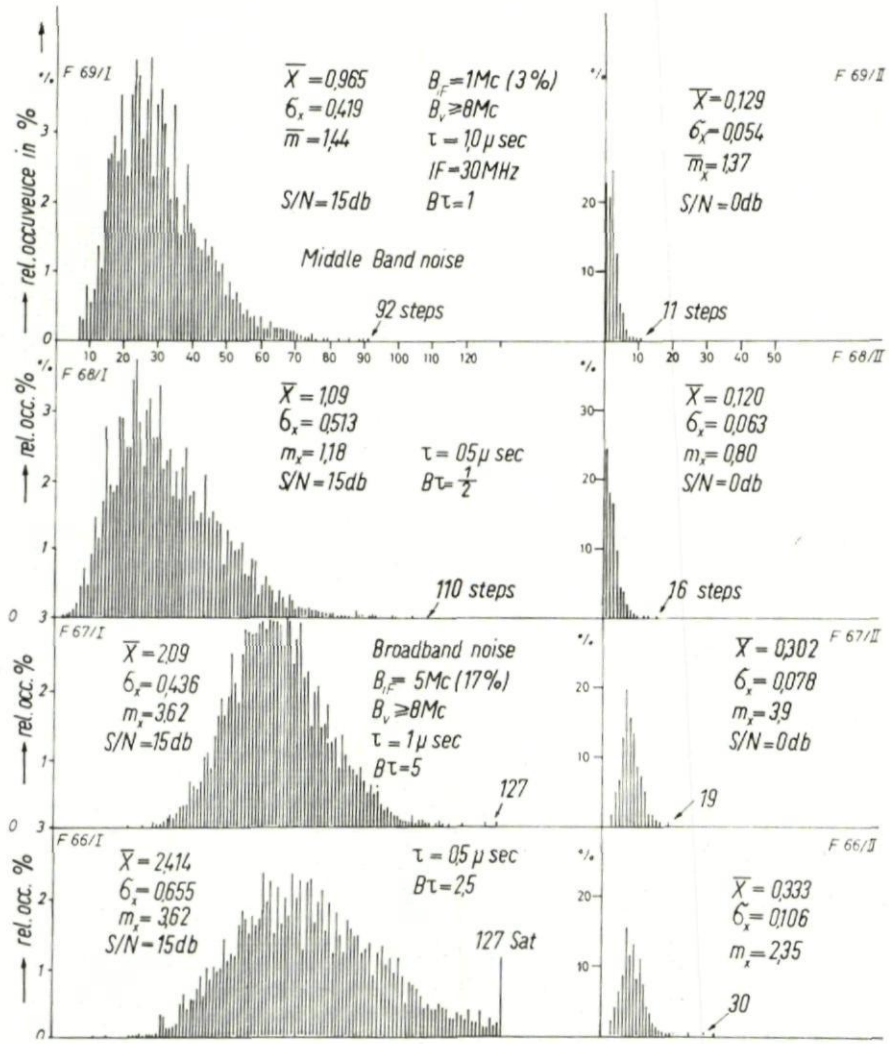


FIG. 8. *Distributions from noise calibration, influence of bandwidth and integration time.*

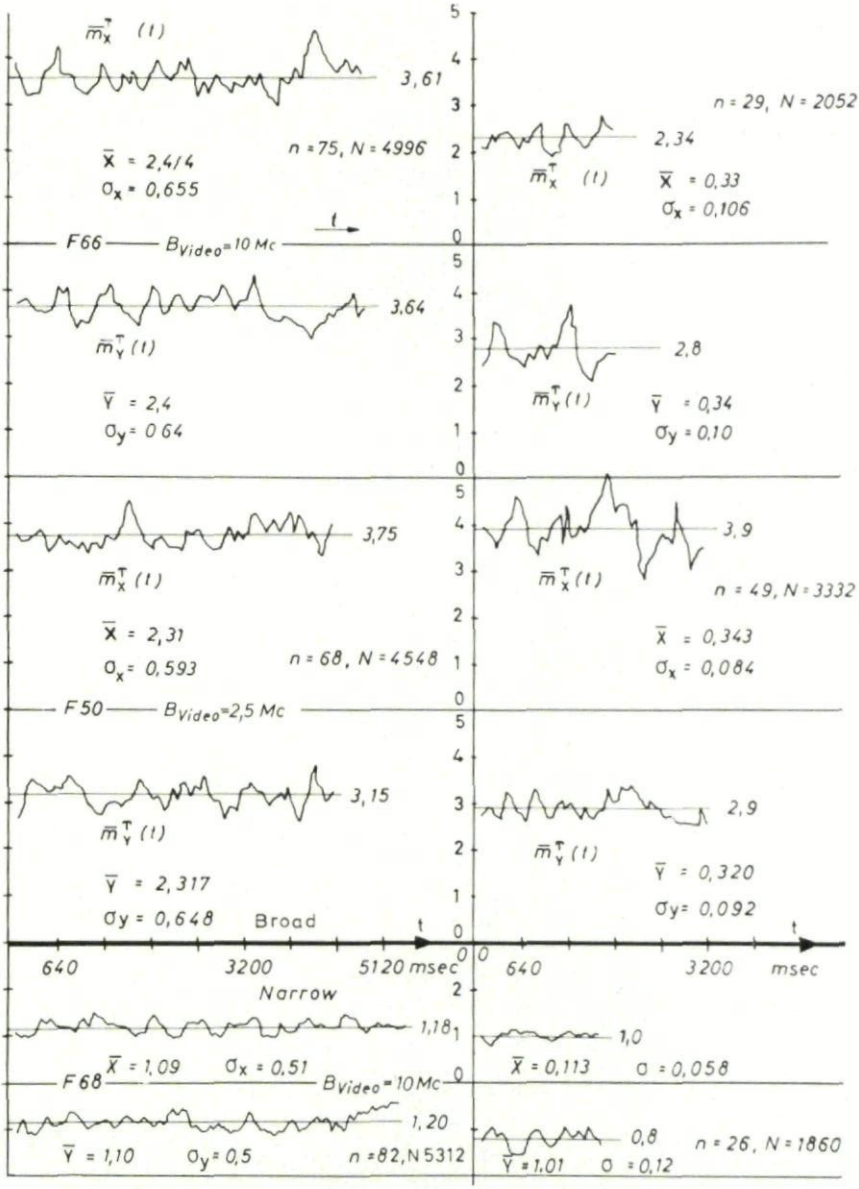


FIG. 9. *Records of noise calibration*

$\overline{m}^{-T}(t)$ *for* $\frac{\text{broadband (5 Mc/s) noise}}{\text{narrow band (1 Mc/s) noise}}$ S/N $\frac{\text{left 15 db, I}}{\text{right 0 db II}}$

$T = 256$ *ms* $x \to \lambda_1$, $y \to \lambda_2$

of video low pass bandwidth, the second form of postdetection integration after linear detector, has been investigated too for the two cases $B_{\text{video}} = \frac{1}{2} B_{\text{IF}}$ and $B_{\text{video}} > 2 B_{\text{IF}}$ for the short range gate 0.5 μs, but no indicative difference could be stated. We are therefore sure, that range gate integration and wide band noise character are first responsible for the lack of probability at low levels, which leads to an increased value of m. Measurements at 0.25 μs range gate are planned. It is possible to calculate the S/N ratio from the DD-functions, if attenuations in db during run and comparative noise measurement are known. A test on the basis of Fig. 8 using noise as a signal, delivers wrong results ($+10 \div 30\%$ error) because of the "bit error," independently of the used moments of the DD, the $\bar{x} =$ (average) or the $\sigma_x =$ (RMS) ratio. Two runs (one for each gate length) with narrow band noise were made using a filter amplifier for constant power density. A test made between corresponding wide and narrow band noise runs (5 and 1 Mc/s bandwidth) shows fair agreement with $\sqrt{5}$ for the average and (RMS) ratio because, although the processes are different in carrier bandwidth, both bandwidths are large with respect to half the PRF. (Sample theorem.)

5.2.

Weather signals have been compared with receiver noise. If errors can be avoided, calibrations can be made on the basis of average or RMS-values. Let β be the IF-channel attenuation used, N representing for the moment the receiver noise level, one gets

$$\left[1 + \frac{S}{N}\right]_{dl} = \left[\frac{(\bar{x} \text{ or } \sigma_x)_{S+N}}{(\bar{x} \text{ or } \sigma_x)_N}\right]^2_{dl} + (\beta_{S+N} - \beta_N) \Rightarrow \left(\frac{S}{N}\right)^{if} \geqslant 1 \qquad (28)$$

Some results on weather noise signals are presented in Figs 10 to 13. Figure 10 . . . shows for the runs F121/I/II and F124/I/II stratosclouds distribution histograms, taken over the whole run, part I being taken before, part II after aeroplane crossing. The values for $\bar{x}$, σ_x ($\bar{y}$, σ_y resp.) are in volts. The averaged Nagakami m-parameter shows variability from 1.5 up to 2.8. The clouds have been almost homogeneous low turbulent icefree stratos in low height (600–825 m). The S/N ratio was not high, on λ_1, $6 + 8$ db, on λ_2, $2 + 3$ db. Analog records showed fair stability within the observation time.

Figure 11 shows the time variations of the short-time averaged m-parameter over the run, in the upper part for the same runs as in Fig. 10, below for two other runs, one in strato clouds (F129/I) the other in rain (F94). Comparison with the corresponding figure from a noise run (Fig. 9) indicates considerable larger $\bar{m}^T(t)$ fluctuations, even if we ignore the low S/N ratio. m-variabilities from 1 to 4 occur within half to one second periods. No significant differences can be stated from a comparison of the clouds before and after the passage of the aircraft.

Run F94 (Fig. 11) shows the same for rain without the influence of receiver noise, the m-fluctuations are on each channel smaller than in the case

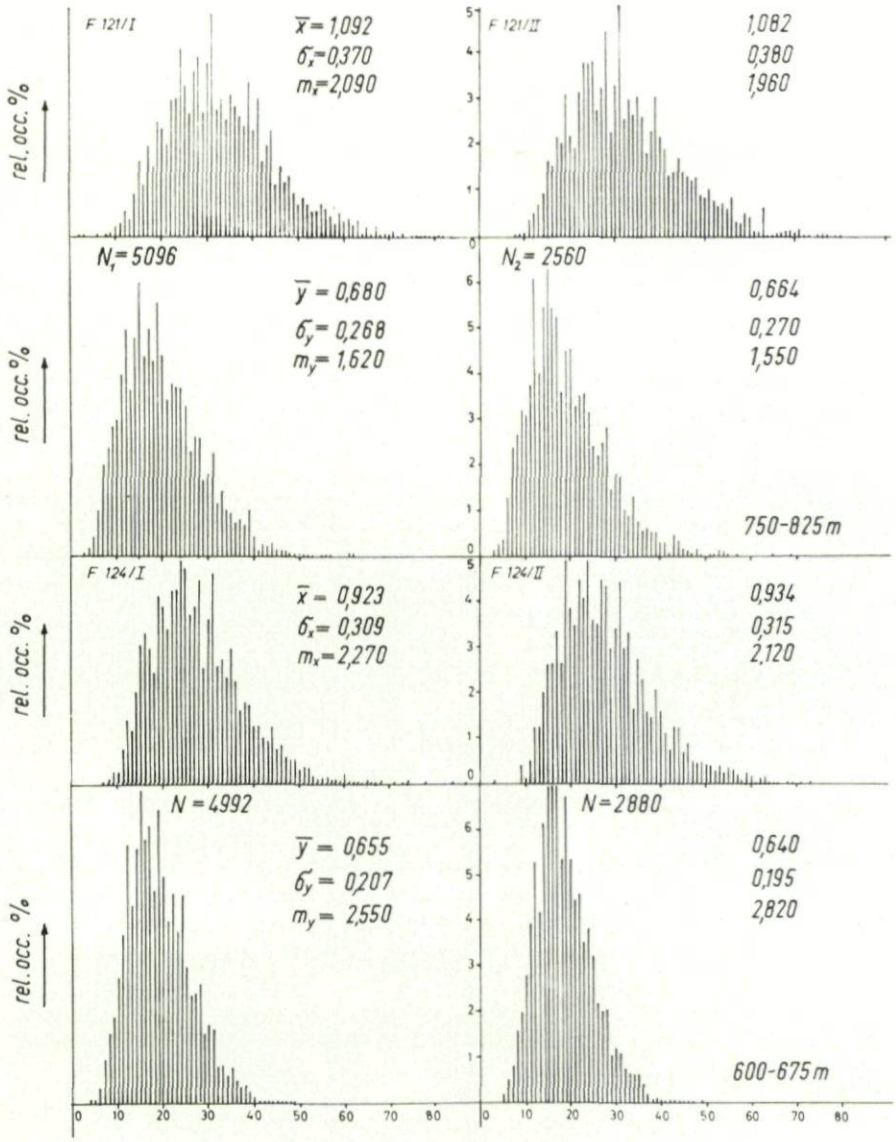


FIG. 10. *Distribution histograms, I before and II after aircraft crossing* ($x \rightarrow \lambda_1 = 0.86$ cm, $y \rightarrow \lambda_2 = 3.2$ cm)

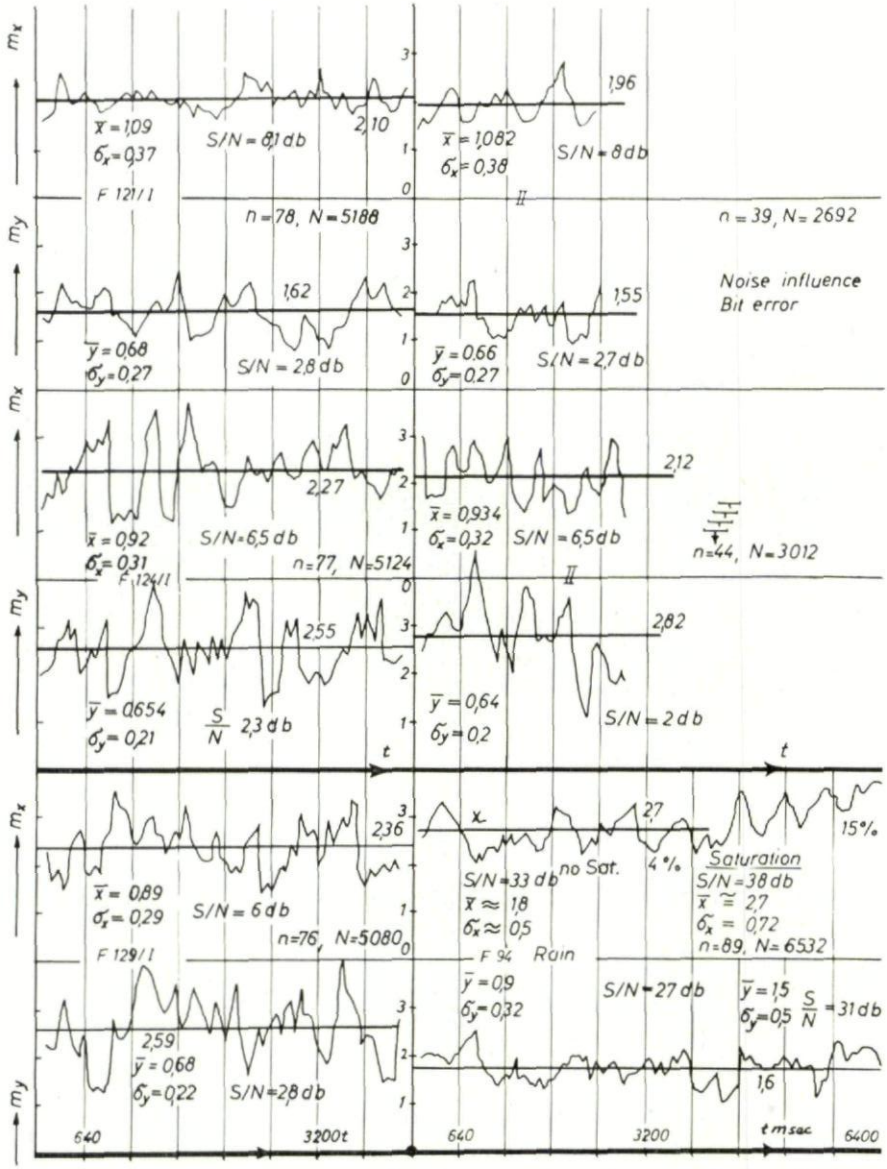
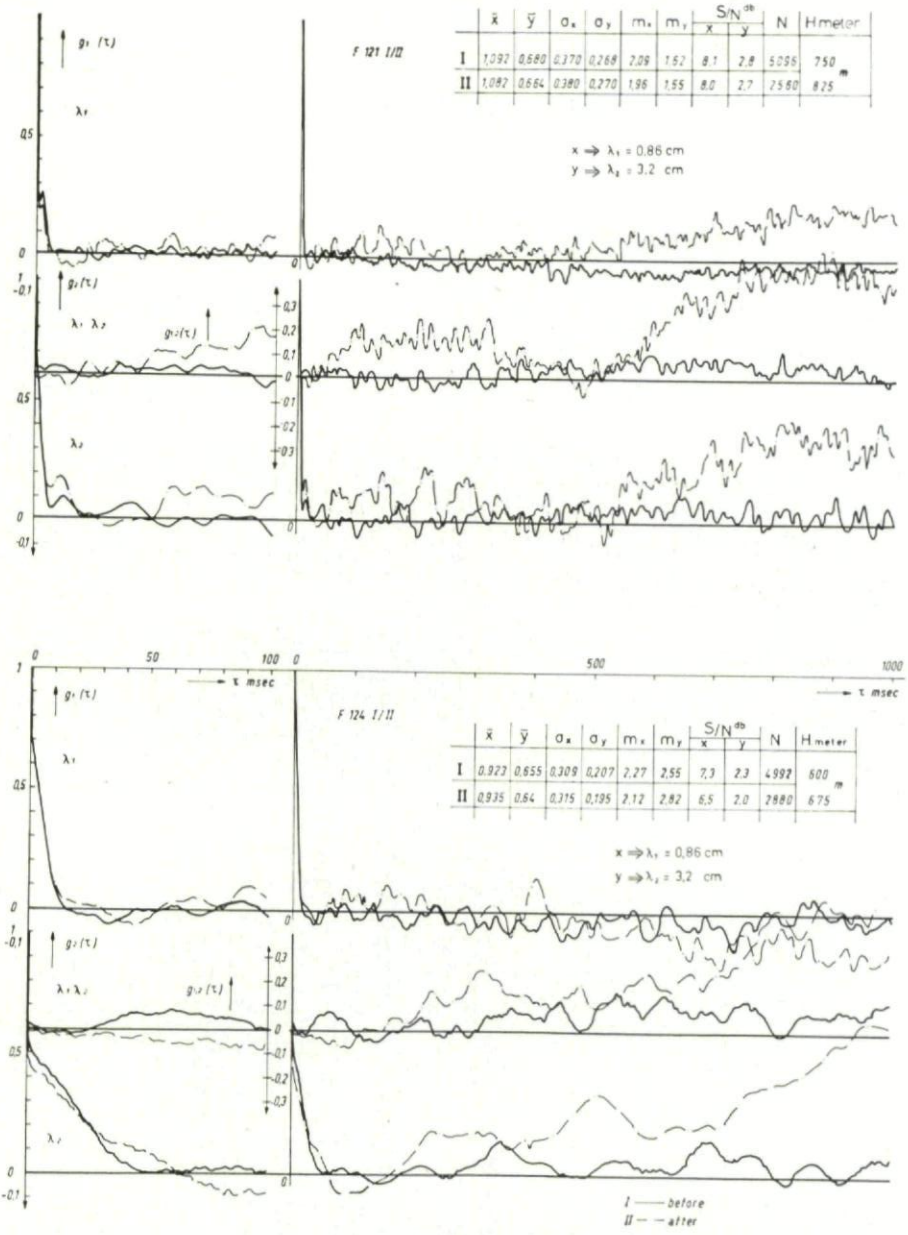


FIG. 11. *Records of signals,* $m^{-T}(t)$ *for* $T = 256$ *ms for strato clouds and rain (F94); I before and II after aircraft crossing,* $x \to \lambda_1 = 0.86$ *cm,* $y \to \lambda_2 = 3.2$ *cm, height* 700 *m*



	$\bar{x}$	$\bar{y}$	σ_x	σ_y	m_x	m_y	S/N db x	S/N db y	N	H meter
I	1,097	0,680	0,370	0,268	2,09	1,62	8,1	2,8	5056	750 m
II	1,082	0,664	0,380	0,270	1,96	1,55	8,0	2,7	7560	825

	$\bar{x}$	$\bar{y}$	σ_x	σ_y	m_x	m_y	S/N db x	S/N db y	N	H meter
I	0,923	0,655	0,309	0,207	2,27	2,55	7,3	2,3	4992	600 m
II	0,935	0,64	0,315	0,195	2,12	2,82	6,5	2,0	2880	675

FIG. 12. *Auto- and cross-correlation functions in strato clouds (I before and II after aircraft crossing)*

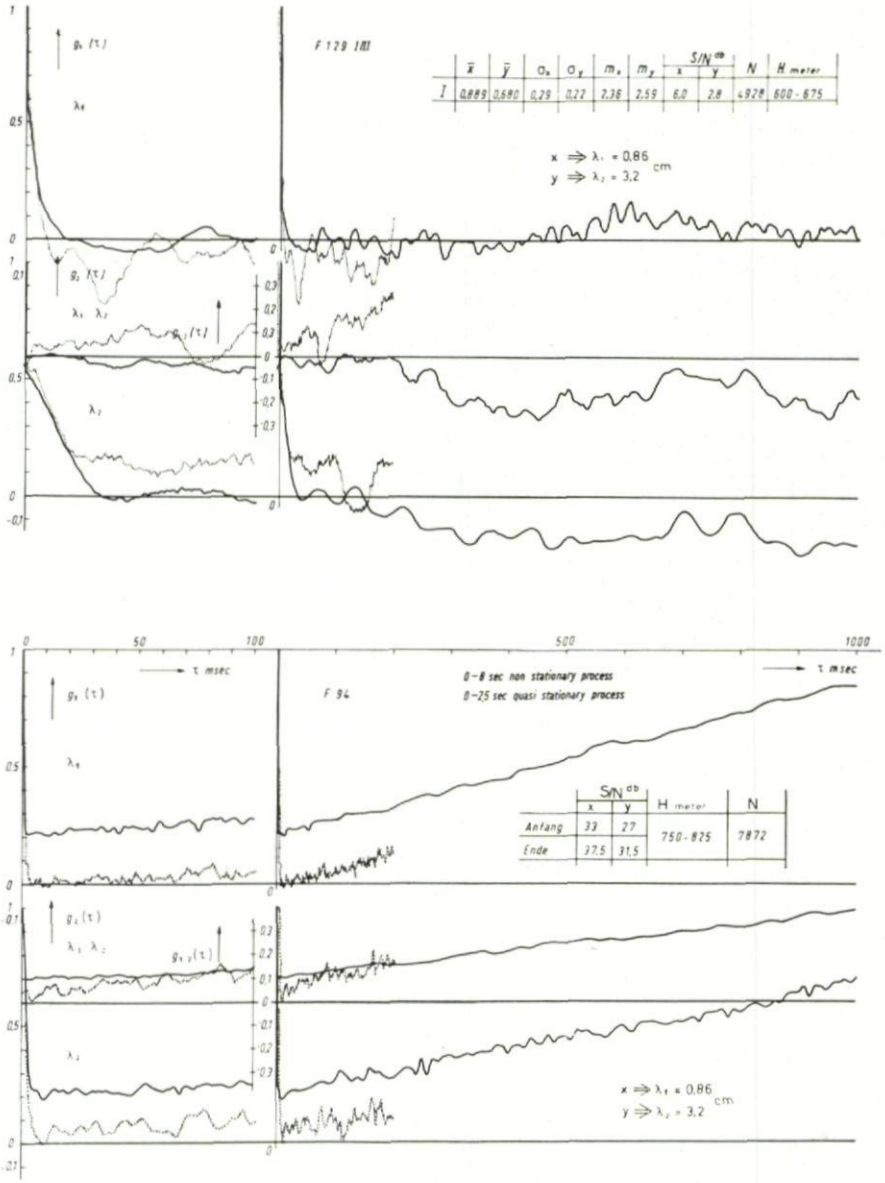


FIG. 13. *Auto- and cross-correlation functions, upper part-strato cloud* (*F*129/1) *and lower part-rain* (*F*94)

of the stratos. On x-channel there is an increase of m beginning with the 3rd second of run, on y-channel the m-value remains fairly constant. The short time averages $\bar{x}$ and $\bar{y}$ both show a considerable increase during the run, which indicates an increase of radar reflectivity by 4 + 6 db. Therefore the signal process was not stable. From this result it follows that only the saturation effect was responsible for the increase in $\bar{m}_x{}^T$. An interpretation for the higher $\bar{m}$-value on the shorter wavelength $\bar{m}_x = 2.7$ compared with $\bar{m}_y = 1.6$ on the longer wavelength could not yet be found. A homogeneous filled pulse volume should deliver the lower $\bar{m}$-value with the x-channel, since pulse volume is smaller for the x-channel. A nonhomogeneous distribution of radar-reflectivity might be the reason for the anomalous m-values.

Figures 12 and 13 show normalized AKF and KKF for the same runs without any correction. On the left side, τ-scale is 1 mm/msc, the right hand time scale is shortened by factor 4. The curves have been drawn directly from the computer. The full line represents the results before cross of aeroplane (I), the dashed line those after the passage of the aircraft.

From a comparison of I with II for each run (121, 124, 129) can be stated in a qualitative sense, that the degree of turbulence inside the pulse volume has been increased after the passage of the aircraft. In F121 the crossing was very accurate, during F124 the plane crossed only nearby and only small side lobe echoes could be observed. On λ_1 channel (sharper beam) there is no indicative variation on F124 but on λ_2 channel (broader beam) there is a clear indication of increased turbulence after aircraft passage. F129 shows qualitatively the same result as F121, the aeroplane crossed the beam more accurately.

To get an orientation about interpretation of fluctuations of the AKF's greater than $\pm$ 0.1 a comparison with corresponding figures from noise runs indicate that fluctuations of AKF more than $\pm 5\%$ are significant. Further, the influence of S/N ratio on the "time to independence" is seen very clearly by looking at the first msec-decay of AKF. But there are some discrepancies, which are not due to equipment errors.

For instance on F121/I or II from the λ_1-AKF the S/N ratio seems to be not more than 1–2 db because of the very rapid decay, however, the measured S/N ratio was 8 db. After correction of S/N ratio a considerable part of decay remains for the first ms due to weather signal. That indicates for the 8.6 mm wavelength that the cloud signal power spectrum was considerably broader than $\frac{1}{2}$ PRF. A look on the video signal on film itself (see for an example Fig. 14) indicates the presence of very fast fluctuations within the gate length, a considerable part of which was doubtless due to receiver noise, but the remaining spectral energy up to the MC range due to cloud-signal. On the 3.2 cm wavelength the measured S/N ratio was only 3 db, from the AKF one would suggest quite the same, which indicates that the weather signal was sufficient narrow-band with only low energy outside the $\frac{1}{2}$ PRF range. The corresponding video signal fast fluctuations cannot be distinguished from receiver noise.

Figure 13 shows the AKF's on both channels for the run F94. The very strong signal with a high S/N ratio was coming from rain falling out from a

strato-cloud. Unfortunately the signal was rising during the run approximately 5 db on both channels, therefore stability conditions were not fulfilled after the first 2.5 s. The instability caused the AKF's to increase monotonously if time averages are taken over the whole film (~8 seconds). The dashed curves show the result, for only the first 2500 ms, up to which the process could be thought of as stationary. The λ_1 AKF falls down to nearly 0.1 within the first ms and remains at low values. This indicates uniquely,

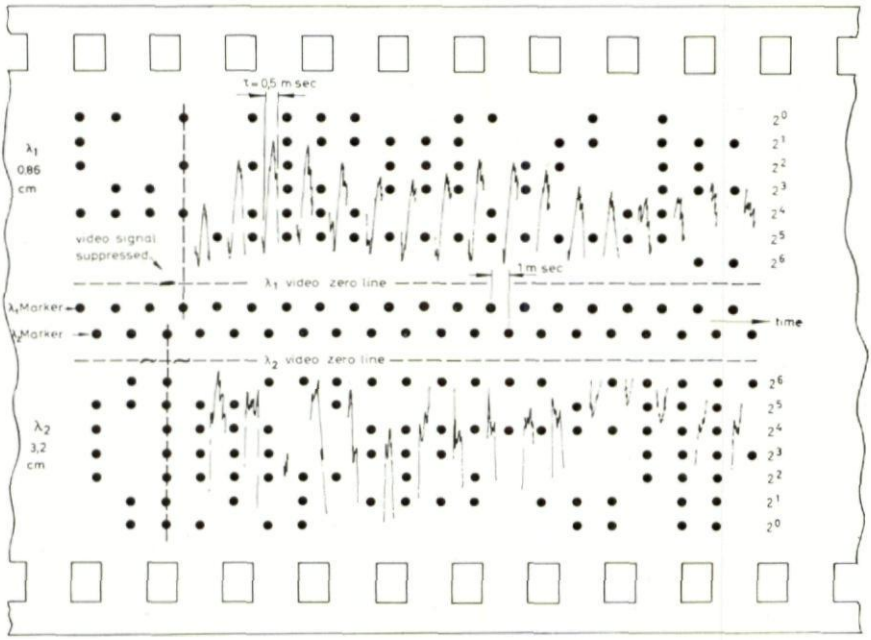


FIG. 14. *Fast fluctuations of range gated video signal (F94 rain S/N better than 30 db).*

because of the absence of noise, that most fluctuation energy is due to relative velocities higher than 2.15 m/s. The λ_2 AKF, which decays down within 4 ms, indicates a considerable part of fluctuation energy due to relative velocities higher than ~8 m/s, because the decay during first ms cannot be due to noise. (It is known, that the max. relative velocity which can be resolved by a radar using wavelength λ and PRF $= f_1$ is given by a relative displacement of half a wavelength per twice the PRF period.

$$V'_{\max} = \frac{\lambda_{1,2} \cdot f_1}{4}$$

This gives for our equipment the above mentioned values 2.15 m/s for λ_1 and 8.0 m/s for λ_2.)

However, Fig. 14, a copied piece of film F94, shows a video trace typical for this run. On both channels we note very fast fluctuations during gate

length (half a microsecond). This indicates the presence of spectral energy up to at least 1 Mc/s doubtless not due to receiver noise. The decay of AKF on both channels may be explained by the "two groups of droplets" model, one consisting of a very high number of cloud droplets, the other a small number of rain droplets falling with their terminal speeds depending on their sizes (2 + more than 8 m/s). For the extremely high frequencies seen in the video trace of Fig. 14 there is no interpretation. Physical velocities in the order of several 1000 m/s do not exist and droplet growing, collisions, and splitting effects are relative slow processes. Further it is very unprobable that the height-average changes so fast with height.

One might ask why the pulse volumes on the two wavelengths have not been made equal. The equipment has been built first to have equal sensitivity in average radar reflectivity together with a good cloud resolution. This was several years ago realisable within a moderate budget frame customed for meteorological investigation only by optimal use of a given antenna at both wavelengths. That led to a sharper beam at 8.6 mm. Technical improvements of the last years would make it possible to use equal beams at both wavelengths, though the realisation of the feed is much more complicated. This means a loss in possible cloud resolution at shorter wavelength. On this point a trade off comes out, because both modes of operations are of quite equal interest: the existing mode with its ability to resolve inhomogeneous distribution of radar reflectivity inside the volume, the mode with equal pulse volumes for more accurate comparisons. Higher transmitter peak power, higher PRF and shorter gate length (0.1 + 0.25 μs) combined together with techniques which make a larger dynamic range available for linear analysis of distributions (up to 70–80 db) and correlations are in any case worth being used in future, because S/N ratios and existing velocity limits will be higher and video averaging effects smaller. If the complete analysis of the gathered material will be finished, results will be made available.

ACKNOWLEDGEMENTS

I am indebted to DVL-Institute of "Flugfunk und Mikrowellen" and to "Deutsche Forschungs Gemeinschaft" (DFG) for financial support of the work, further to the flight crew, the pilot S. Loose, M. Reinhardt of the DVL-Institute of "Physik der Atmosphäre" (former "Flugwissenschaftliche Forschungsanstalt München" FFM), to Dr. Splechtna for the development of digital equipment and computer calculations, to H. Borchardt for his meteorological assistance and to many of my institutes colleagues for useful discussions and operational work immediately with the radar.

DISCUSSION

N. J. Smith: When you wish to compare on two frequencies are you concerned with the different volume of cloud illuminated due to difference of beam width. In many similar experiments we have been carrying out on Q-Band we have observed fluctuations, as you said, lasting over a few seconds

and we are attempting to find out whether it is a time variation that causes these fluctuations or a special fluctuation where the cloud, or in our case the rain, drifts through the illuminated volume. When you are averaging some sixteen times the volume in the case of *S*-Band compared with *Q*-Band is it fair to compare the results directly as appropriate to one volume?

W. FOGY: We cannot distinguish between time and special fluctuations because an instantaneous amplitude represents a space average. In planning our equipment we decided to have different volumes at *Q* and *X* bands in order to concentrate the *Q*-Band power so as to give an adequate signal to noise ratio. There are advantages and disadvantages in choosing different volumes at the two frequencies.

REFERENCES

1. KERR, D. E., *Propagation of Short Radio Waves*, MIT series, vol. 1.
2. SCOLNIK, M. I., *Introduction to Radar Systems*.
3. SIEGERT, A. J. F. and MARTIN, F. W., Fluctuations in the return signals from random scatterers (Window, rain, sea, Echo), *Radiation Lab. Rep.* **773,** Jan. 1964.
4. WALLACE, P. R., Interpretation of the fluctuating echo from randomly distributed scatterers—Part 2, *McGill Univ., Montreal, Res. Rep.* **MW-6,** Dec. 1951.
5. ATLAS, D. and BARTNOFF, S., Cloud visibility, radar reflectivity and drop-size distribution. *J. of Meteorology*, **10,** April 1953.
   ATLAS, D., The estimation of cloud parameters by radar, *J. of Meteorology*, **11,** Aug. 1954.
6. AUSTIN, P. M., A study of amplitude distribution function for radar echoes from precipitation, *MIT, Tech. Rep.* **17,** May 1952.
7. FLEISHER, A., The information contained in weather noise, *MIT, Tech. Rep.* **22,** Jan. 1953.
8. HILST, G. R., Analysis of the audio-frequency fluctuations in radar storm echoes, *MIT, Tech. Rep.* **9,** Nov. 1949.
9. HITSCHFELD, W. and DENNIS, A. S. Measurement and calculation of fluctuations in radar echoes from snow, *McGill Univ., Montreal, Sci. Rep.* **MW-23,** July 1956.
10. MARSHALL, J. S. and HITSCHFELD, W., Interpretation of the fluctuating echo from randomly distributed scatterers—Part 1, *McGill Univ., Montreal, Res. Rep.* **MW-4,** October 1951.
11. ROGERS, R. R., Radar measurement of gustiness, *MIT, Res. Rep.* **29,** Feb. 1957.
    ROGERS, R. R. and PILIÉ, R. J., Radar measurements of drop-size distribution, *J. Atmos. Sci.* **10,** Nov. 1962.
12. STONE, M. L. and FLEISHER, A., The measurement of weather noise, *MIT, Res. Rep.* **26,** May 1956.
13. RUTKOWSKI, W. and FLEISHER, A., R-Meter: an instrument for measuring gustiness, *MIT, Res. Rep.* **24,** Oct. 1955.
14. LHERMITTE, R. M., *Echo Fluctuations Study with Zenith Pointing Radar*, Observatoire de Magny-les-Hameaux, Seine et Oise, France.
    LHERMITTE, R. M., Weather echoes in Doppler and conventional radars, AF Cambridge Research Laboratories, *Proc. 10th Weather Radar Conference.*
    R. M. LHERMITTE, R. M., New developments of the echo fluctuations theory and measurements, *Proc. 9th Weather Radar Conference.*
15. PROBERT-JONES, J. R., *The Radar Equation in Meteorology*, Department of Meteorology, Imp. College, London.

16. KLEEGREWE, C., Bau eines Wolkenradargerätes zur gleichzeitigen Messung bei 3.2 cm und 0.86 cm Wellenlänge, *DVL-Rep.* **96,** 1959.
17. BORCHARDT, H., Physikalisch-technische Grundlagen der meteorologischen Anwendung von Radar nach Erfahrungen mit der Wetterradaranlage des Instituts für Mikrowellen in der Deutschen Versuchsanstalt für Luftfahrt e.V., Mühlheim (Ruhr), *DVL-Bericht*, **109,** Mar. 1960.
BORCHARDT, H., Wolkenbeobachtungen mit einem doppelwelligen Radargerät, *Beiträge zur Physik der Atmosphäre*, **35,** 1/2, 1962.
18. STACKPOLE, J. D., Radar measurement of the spectra of turbulence in the free atmosphere, *MIT, Res. Rep.* **32,** 1959.
19. WILLIAMS, E. L., The pulse integrator, part A: description of the instrument and its circuitry, *MIT, Tech. Rep.* **8,** Aug. 1949.
WILLIAMS, E. L., The radar signal spectrograph, part B: description of the instrument and its circuitry, *MIT* Tech, *Rep.* **9,** Jan. 1950.
20. NICOL, R. W. E., The fading of radio waves of medium and high frequencies, *Proc. Instn. Elect. Engrs.*, Pt. III, **96,** paper no. 891.
21. —— *Meteorolog. Res. Rev.*, **3,** 14, July 1957.
22. ZUHRT, H., Die Summenhäufigkeitskurven der exzentrischen Rayleigh-Verteilung und ihre Anwendung auf Ausbreitungsmessungen, *Archiv der elektrischen Übertragung*, Siemens & Halske A. G.
23. NAGAKAMI, M., The *m*-Distribution—a general formula of intensity distribution of rapid fading, *Statistical Methods in Radio Wave Propagation*, Pergamon, 1960.
24. VOGEL, M., Die Korrelation zwischen den Streuquerschnitten eines gleichzeitig bei zwei verschiedenen Frequenzen beobachteten Flugzeuges, *Internal DVL-Report.*

CHAPTER 9

MOVING TARGET INDICATION A SURVEY OF DEVELOPMENTS SINCE 1948

P. BRADSELL
Decca Radar Ltd., Isle of Wight Division, England

1. INTRODUCTION

The art and science of moving target indication (m.t.i.) has advanced considerably over the last fifteen years, and it is the aim of this paper to survey the principal developments in theory and technique which have made this progress possible.

Moving targets had been detected in radar experiments using c.w. in the 1930's; however, c.w techniques are outside the scope of the present paper. By 1948, the master oscillator-power amplifier (m.o.p.a) system shown in Fig. 1 and the coho-stalo system of Fig. 2 were well established, and non-coherent m.t.i. had been used in airborne radar systems.

At frequencies below 1000 Mc/s, where r.f. power amplifiers were available, m.o.p.a. radar transmitters continued to be used, but for *L* and *S*-band transmitters the magnetron was preferred, being a cheap, compact source of r.f. power requiring moderate supply voltages. Because the magnetron transmits with random initial phase, m.t.i. systems were mostly of the

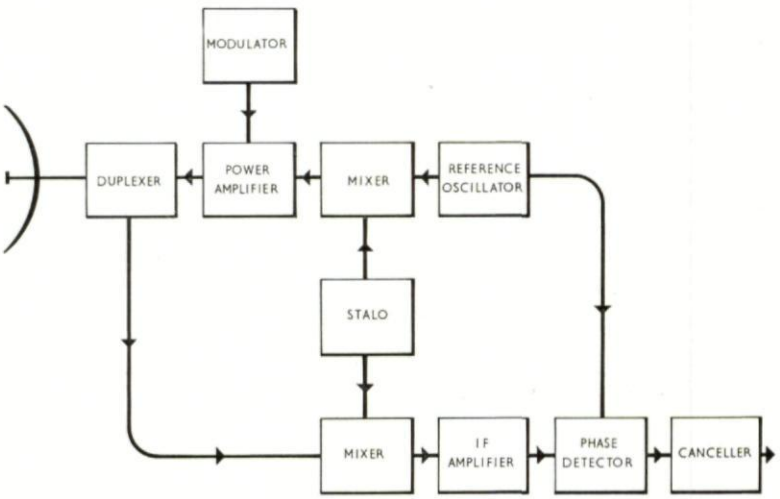


FIG. 1. *MOPA m.t.i. system.*

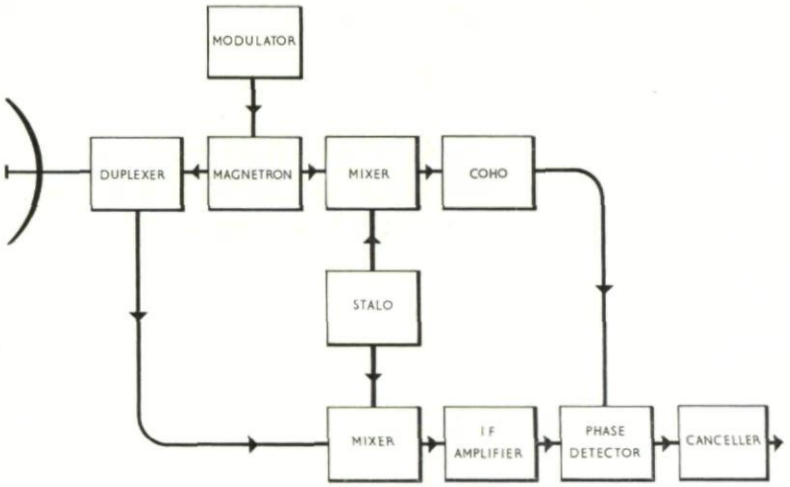


FIG. 2. *Coho-stalo m.t.i. system.*

cohostalo type. There have been striking improvements in components for these systems, and extensions to the capability of the systems following fuller understanding of the theory of m.t.i. Increasing complexity has demanded greater reliability, and high-lighted the need for measurement of performance in the field.

These topics are discussed below, followed by a look at the new generation of m.t.i. equipment which is coming into use.

The frequency bands in which coherent m.t.i. has been used range from

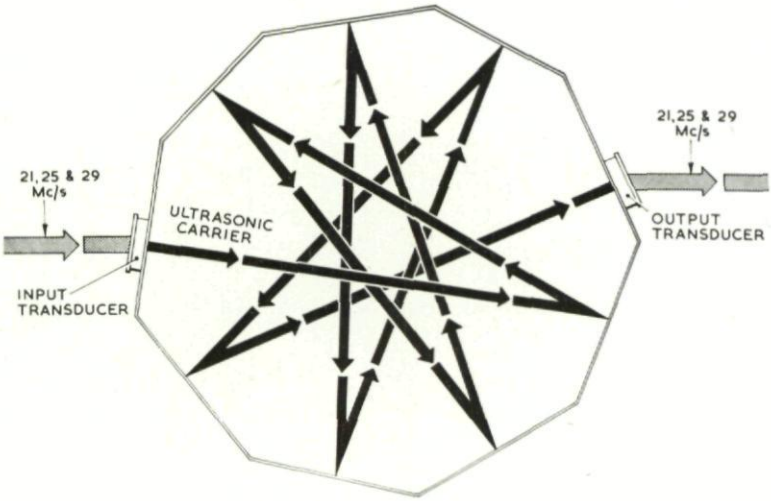


FIG. 3. *Quartz ultrasonic delay line.*

v.h.f. (400 Mc/s) to *S*-band (3000 Mc/s). Successful m.t.i. at 10,000 Mc/s and above has been mainly of the non-coherent type, but nowadays coherent techniques are possible up to the highest radio frequencies.

2. COMPONENTS FROM M.T.I.

In this section an attempt has been made to treat the subjects in order of importance, as seen by the writer.

2.1. *Delay Lines*

One or more ultra-sonic delay lines are an essential part of most permanent echo cancellation systems. Signals are stored in the delay line for a time equal to the interpulse period of the radar, so that they can be subtracted from the signals following the next transmitted pulse. The delay channel must reproduce the signals without distortion so that the residue after subtraction is a minimum.

The first delay lines used in m.t.i. were liquid filled (either water or mercury) with quartz transducers.[1] The attenuation in a liquid increases as the square of the frequency, so that the choice of carrier frequency is a compromise between attenuation and bandwidth. The design was complicated by the need to correct for the attenuation slope across the pass band. The longer mercury lines (over a millisecond delay) had to be driven with several watts of carrier power and followed by a carefully shielded low noise amplifier.

An attenuation exceeding 90 db for 2-3 Mc/s bandwidth was typical for a particular 1.25 ms line. The final stages of manufacture were extremely critical, and the product neither repeatable nor constant with time. It is fair to say that some manufacturers conquered these difficulties and produced satisfactory delay elements; however, new designs of equipment now use quartz or fused silica lines[2].

In these, the delay medium takes the form of a polygon of silica (Fig. 3), to which are bonded quartz crystal transducers. Multiple folding of the delay path makes the device compact, and the performance is stable and reliable. Typical attenuations range from 40-60 db, carrier frequencies from 10-60 Mc/s. The bandwidth may be 30-50% of the carrier frequency so that where the bandwidth of the modulated carrier is small, the line can be multiplexed to give a number of independent delay channels. Alternatively matched groups of delay lines may be mounted in a common temperature controlled enclosure.

Production of large solid delay lines is dependent on the availability of quartz or silica plates of fine optical quality and suitable size; this was at one time a matter of extreme difficulty but delays of 4-5 m are now possible. Intensive development has enabled spurious signals (i.e. signals arriving via incorrect delay paths) to be brought down to a level −40 db to −50 db with respect to the wanted carrier. Production is a highly specialised process, giving delay lines of excellent performance at a competitive price. Other applications are incremental delays (100–500 μs) staggered p.r.f. systems and test delay lines (10–60 μs) which generate coherent test signals by

multiple reflection of the coho lock pulse (at i.f.). The delayed signals are fed to the input of the coherent receiver, and appear at the output as artificial fixed targets.

The temperature co-efficient of silica (approximately -80 in $10^6/°C$ is less than that of mercury, and for many purposes temperature control is unnecessary, the radar p.r.f. being locked to the delay line. Where several delay elements are associated with one p.r.f. (e.g. in multiple-beam radars), individual delays can be matched to within 5 parts in 10.[6]

Alternative delay media such as metals, ceramics and glasses are being investigated, without a serious challenge to quartz or silica, at least for delays exceeding a few hundred microseconds.

Ceramic transducers, of materials such as barium titanate[3] offer lower losses in conversion of electrical to ultra-sonic energy and vice versa, but are not yet in general use. It has recently been discovered that semi-conductor junction layers subjected to alternating electric fields can launch ultrasonic waves; this effect promises to extend the frequency range of ultrasonics into the v.h.f. and u.h.f. bands.

The quantum aspects of ultra-sonic energy are also being studied.[4] Amplification of ultrasonic waves in a medium has been demonstrated; the principle is similar to that of a maser, the pump power being supplied by a beam of light.

2.2. *Stalos*

A stalo is a stable local oscillator. Present day applications require spurious f.m. deviations to be smaller than 10–100 c/s, often under severe vibration conditions. Remote tuning is also needed to follow long term variations in transmitter frequency.

The first microwave stalos were klystrons.[5] The klystron frequency is sensitive to microphony, external magnetic fields and power supply ripple so that even with carefully smoothed supplies additional frequency stabilisation was necessary. A simple method was direct coupling of the klystron cavity to a high Q resonator such as an echo box, the klystron frequency then being determined by the tuning of the stabilising cavity.

Improved stability was obtained with feedback systems as described by R.V. Pound[6,7] in which the klystron frequency is controlled by the output of a microwave cavity discriminator. Stalos of this type were produced at *S* and *X*-band, but were complex and difficult to tune over a wide band.

Triodes had been used as local oscillators in *L*-band m.t.i. systems, and it was shown that, given careful isolation from vibration and smooth supplies, the triode stalo was practicable at *S*-band, without special stabilisation.[8] The valves used range from the war-time lighthouse tubes, through later glass-envelope disc seal triodes to the latest ceramic envelope valves.

These valves were not developed specifically for stalo service, but improvements in design and manufacture have resulted in better stability, reliability, ruggedness and lower power consumption.

Although some manufacturers offer varactor tuned oscillators, the majority of triode stalos are mechanically tuned. A typical modern cavity oscillator

covers a frequency band of 20–30% with a single manual adjustment, a motor driven fine tuner giving automatic frequency control.

A well designed crystal oscillator followed by a chain of frequency multipliers can generate microwave power with better frequency stability than a self-excited oscillator[9,10] (less than 1 c/s deviation at *S*-band, and less than 10 c/s at *X*-band). To achieve this stability the crystal and mount must be carefully isolated from vibration. The oscillator valve or transistor must have a frequency capability well in excess of the frequency of oscillation to minimise phase shift, power supplies including heaters to oscillators and multipliers must be d.c. with very low ripple.

In order to ensure a good signal to noise ratio, the oscillator should be strongly excited. This is in contrast to the standard practice in crystal oscillators of good long term stability which may have inferior performance when measured over millisecond time intervals.

Because of the size, power consumption and complexity, crystal multiplier stalos have found wider application in test equipment than in field service radars.

With the development of efficient varactor frequency multipliers,[11] the "solid state stalo" is now a practicable device. A crystal controlled transistor oscillator drives transistor frequency multipliers giving a watt or more at v.h.f. Varactor stages then multiply the frequency up to the microwave region. This type of stalo will give some tens of milliwatts with good stability in a severe vibration environment.

Disadvantages of the crystal oscillator multiplier are the small tuning range (normally a fraction of 1%), and the presence of spurious harmonic lines in the output. These difficulties may be overcome to some extent by phase locking a klystron oscillator to the crystal harmonic frequency.[12] The klystron output then contains only the desired harmonic, other frequencies being strongly attenuated. The klystron may be locked to any one of a "comb" of frequencies and can then be tuned over a limited range (2–4 Mc/s) by altering the frequency of the i.f. reference oscillator in the phase lock loop.

Certain applications demand rapid selection of frequency over a wide band, together with good short term stability. A voltage tunable magnetron or backward wave oscillator has been coupled to a long microwave transmission line[13] so that oscillation takes place at discrete frequencies. The spacing between stable frequencies is governed by the length of transmission line; the frequencies are selected by a programmed variation of the oscillator beam voltage.

3. *Transistors*

Cheap silicon high-frequency transistors have become available in the last few years for all video and carrier circuits in m.t.i.

Size, weight, power consumption and heat dissipation have all been reduced, and complex equipment will now function for months on end without component failures, and with only the simplest maintenance (Fig. 4).

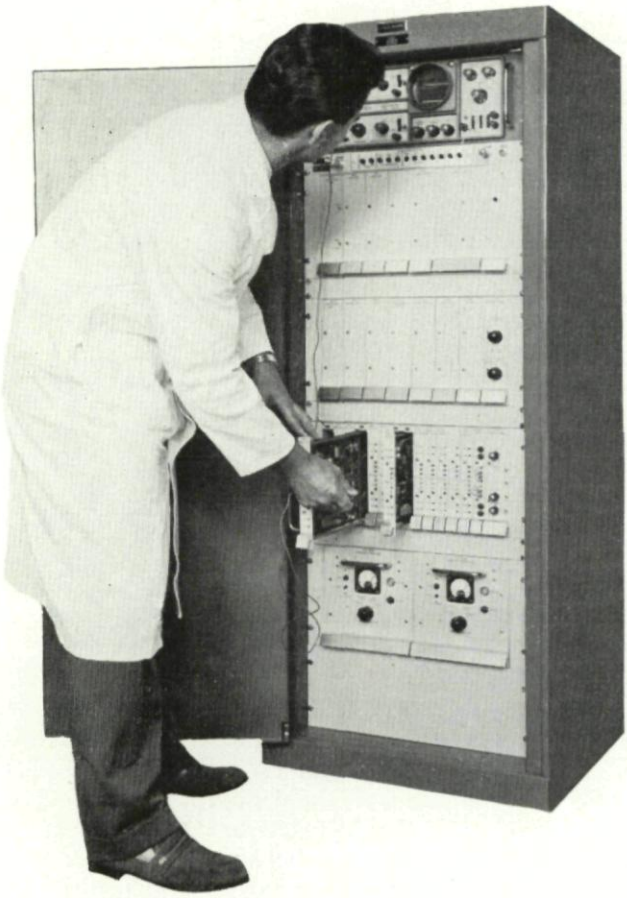


FIG. 4. *Decca transistorized m.t.i. rack.*

2.4. *Phase Detectors*

The conventional balanced phase detector,[14] with either thermionic or semi-conductor diodes is still widely used.

However, even with optimum adjustment, the slope of the voltage/phase characteristic varies over the cycle in the ratio $1 : \sqrt{2}$.

The 6BN6 gated beam tube phase detector[15] consists of a tetrode with two grids controlling the same electron stream.

I.F. signals and coho reference voltage are fed at high level (20-30 volts) to the control grids so that the electron beam is modulated by a square wave. The on/off ratio of this square wave and hence the mean anode current of the 6BN6 depends linearly on the phase difference between the signals at the input grids.

Unfortunately, a semi-conductor equivalent of this device is not yet available.

2.5. *Modulators*

A typical specification for the transmitter modulator of an m.t.i. radar calls for a time jitter less than one hundredth of the transmitter pulse length.

Spark-gap modulators and mercury vapour thyratrons will not meet this specification (though some mercury-pool switches can). During the period under consideration, the design of hydrogen thyratrons has improved so that time jitter is no longer a restricting factor in the design. This is due to better control of the gas pressure within the valve, and the development of tetrode thyratrons. In the latter, the discharge is first established between the inner grid and cathode, and transferred to the anode when the outer or control grid is triggered.

Hydrogen thryatrons are now available which can drive the largest magnetrons, and are beginning to replace damping diodes in the larger line-type modulators.

Applications involving extremely precise timing (1 to 10 ns) have continued to use hard valves as switches, and these are again coming into use for keying the extremely high voltages in modern klystron transmitters.

3. IMPROVEMENTS IN TECHNIQUE

Recognition of the m.t.i. system as a frequency filter has led to important improvements in equipment performance. The well known velocity response of a single canceller can be drawn against a horizontal axis of frequency, the zeroes (blind speeds) corresponding to harmonics of the radar p.r.f.

The m.t.i. canceller is in fact a filter possessing a periodic transmission characteristic having zeroes which correspond to individual lines of the spectrum of the radar pulse (Fig. 5).

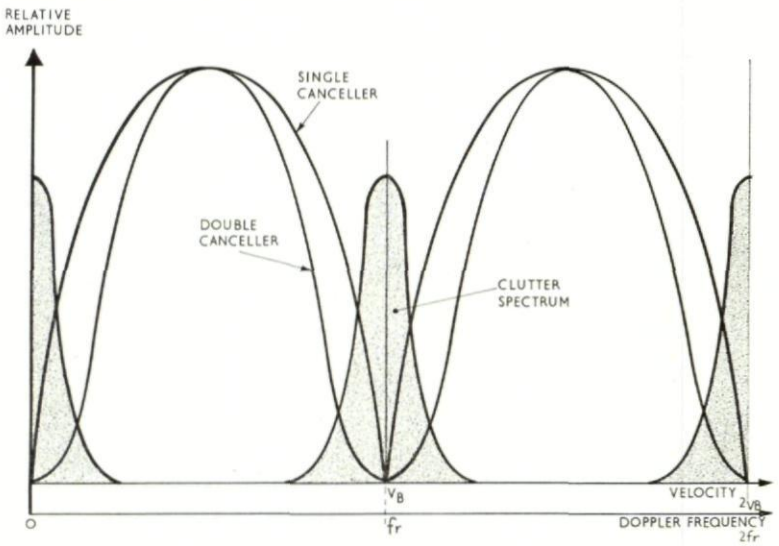


FIG. 5. *M.T.I. velocity responses.*

The cancellation obtainable with a scanning m.t.i. radar is known to be poorer than that of the same equipment with the aerial stationary, to an extent which depends on the number of pulses radiated per beamwidth of aerial scan. An analysis of this effect on a statistical basis is given by Ridenour.[16] In the frequency domain, the line spectrum received by a stationary aerial is modified under scanning conditions, each spectral line being broadened to a shape which is the Fourier transform of the envelope of the pulses received from an isolated fixed target. This envelope is the two-way horizontal polar diagram of the aerial, and its time scale is determined by the aerial rotation rate. The same result follows by summing all the frequency shifted components resulting from transmission by one element of a rotating horizontal array and reception by another element. The form of the spectrum is the auto-correlation function of the aperture distribution, and the maximum frequency spread is equal to the doppler shift corresponding to the speed of the fastest moving elements at the ends of the array.

The effect of scanning is thus to spread the line spectrum of fixed echoes so that the energy lies at frequencies where the filter response of the m.t.i. is not zero; the output of the canceller will contain an uncancelled residue of clutter.

3.1. *Double Cancellation*

Double or cascaded cancellers[17] came into use just over ten years ago. Two similar delay and subtraction circuits are arranged in cascade, the overall frequency response being the product of the frequency responses of the two individual cancellers. In particular, the linear variation with frequency of the single canceller response near zero or a blind speed, becomes a square law variation in the double canceller. The response of the system to frequencies contained in the clutter spectrum is thus reduced, resulting in more effective cancellation of clutter. In a medium range search radar with 10 pulses per beamwidth the improvement due to double cancellation is 10–12 db.

The use of double cancellation would appear to increase the width of the blind speed gap. However, because the clutter residue is at a lower level after double cancellation the system noise will be set to a lower level compared with a single canceller. In consequence the blind speed widths will be very similar if measured at a fixed level with respect to noise in the two cases.

There are no fundamental engineering problems associated with the use of double cancellation; two independent, matched, delay channels must be provided, together with the necessary signal handling circuits.

3.2. *Velocity Shaping*

The technique of optimising the response of an m.t.i. system by the use of a video feedback is known as velocity shaping.[18] The properties of cancellers with feedback have been studied by the use of the Z-transform, which describes the complex frequency response of a periodic filter in terms of a finite number of poles and zeros. The results for a double canceller show that two

feedback paths (Fig. 6) are sufficient to control the blind speed width and the degree of peaking of the velocity response. The effect of feedback is degenerative below the cut-off frequency, giving better clutter rejection than a canceller without feedback, and re-generative above the cut-off frequency giving more uniform response for moving targets.

Engineering difficulties encountered in developing a feedback system are firstly that additional short delay elements must be inserted in the delayed channels of the canceller to compensate for unwanted delays in the undelayed

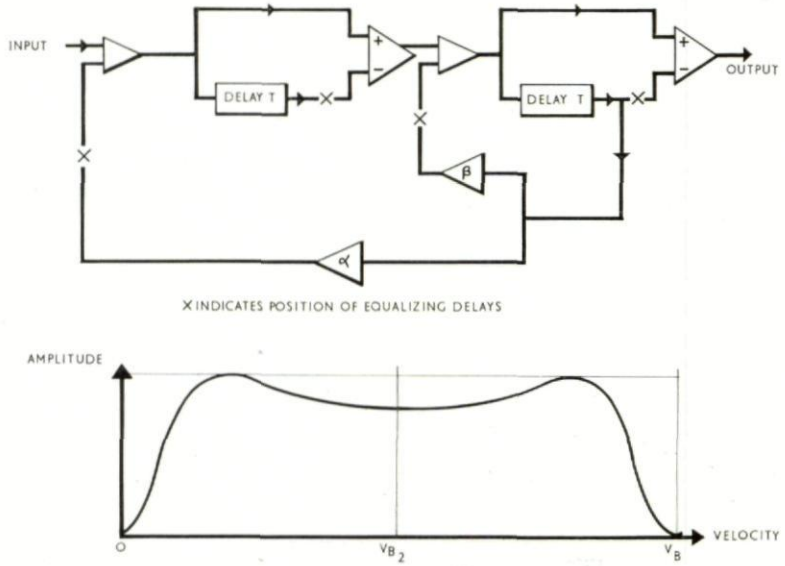


FIG. 6. *Double canceller with feedback.*

channels and in the feedback paths. Secondly, because the signals are recirculated through the delay line by the feedback, a larger bandwidth is necessary for a given degree of pulse distortion.

A useful way of application for feedback m.t.i. appears to be in search radars with 5 to 10 pulses per beamwidth, where double cancellation alone would give insufficient clutter rejection.

3.3. *Staggered p.r.f.*

Another important m.t.i. concept[19] developed in the early 1950's is staggered p.r.f.[1] The non-zero blind speeds for an m.t.i. are dependent on the interpulse spacing; a target moving one half wavelength between pulses will be lost from view. If the pulse spacing is staggered i.e. made non-uniform (Fig. 7) then the blind speed gaps will be partially filled in. The velocity response characteristic (Fig. 8) will be free from zeroes up to a speed, which is the lowest common multiple of the component blind speeds corresponding to the various interpulse intervals.

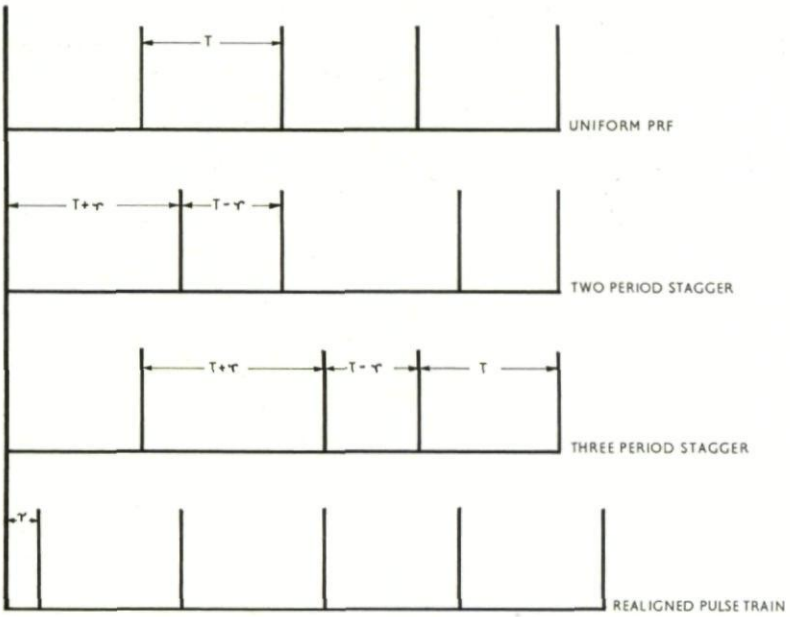


FIG. 7. *Staggered p.r.f. pulses.*

If the ratio of periods is very close say 29:30 (Fig. 9a) then the blind speed is increased by a large factor, in this case 30 times. On the other hand the two component first-order blind speeds are separated by only one part in 30 so that the response falls nearly to zero after the first maximum. The choice of stagger ratio is clearly a compromise between gap free coverage at high target speeds and deep minima at low speeds. The final selection can only be made by examining computed responses for various ratios.

The staggered p.r.f. response for a single cancellation system can be computed by taking the root mean square of the component velocity responses, but this method gives inaccurate results for a cascaded or velocity-shaped canceller in which the output is a function of two or more preceding input pulse amplitudes. An exact analysis for a general form of canceller is given in the Appendix. A recent application of this analysis was in the design of

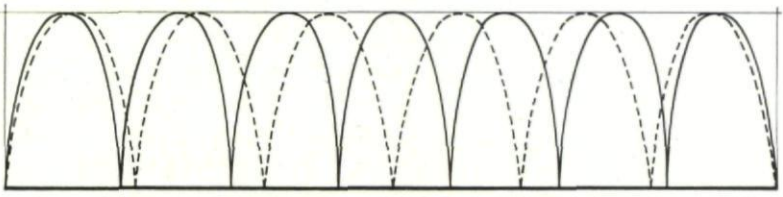


FIG. 8. *Velocity responses, p.r.f. ratio* 6:7.

the Decca AR-1 air surveillance radar, a computer was programmed to select the closest stagger ratio, either two or three-period for which there were no minima below −6 db with respect to the average. The ratio selected was 7 : 8 : 9 (Fig. 9b) giving a first blind speed of 560 knots for a *S*-band radar with mean p.r.f. 700 p.p.s. The use of two radio frequencies in diversity gives a further staggering effect which fills in the 560 knot blind gap on the equipment described.

In a staggered p.r.f. system, the p.r.f. is usually generated with uniform timing and the staggered transmitter trigger derived from this by additional

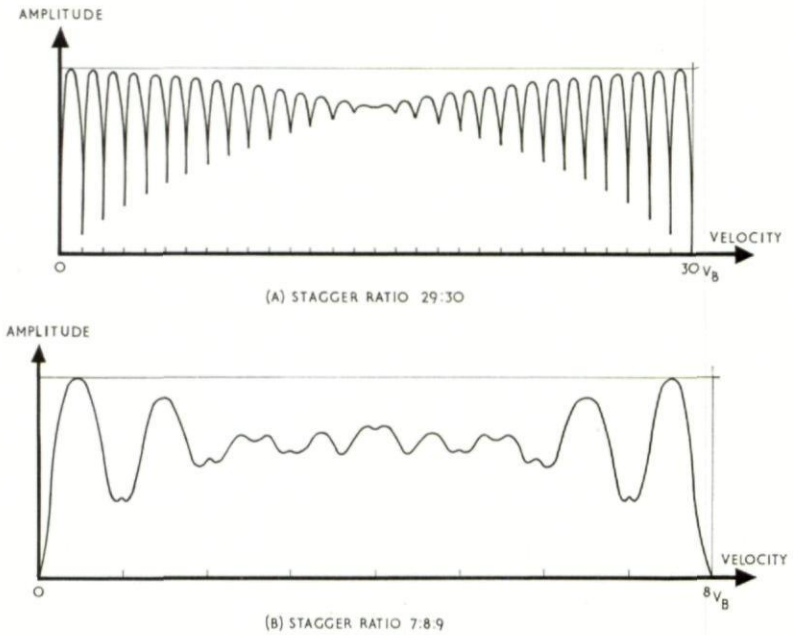
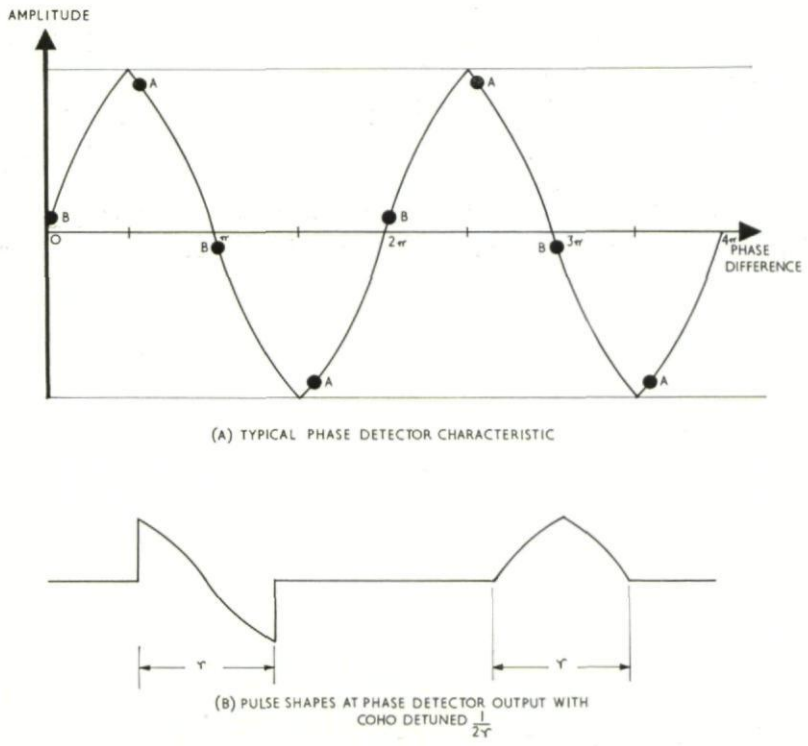


FIG. 9. *Staggered p.r.f. velocity responses.*

time delay in selected periods. The incremental delay line is similar in construction to the main delay lines described in Section 2.1. The same delay is used to re-align the received signals to uniform spacing before cancellation. A similar result has been achieved by a simpler way by switching between two main delay lines; in this method the display is triggered at staggered p.r.f. which may be a disadvantage.

3.4. *Coherent Low i.f.*

In addition to blind speeds, blind phases are present in coherent m.t.i. A moving target at optimum speed occasionally returns a signal so phased with respect to the reference oscillation that the phase detector output is zero (Fig. 10a). This condition gives rise to deep fading in a small proportion of targets, and to pulses missing from the displayed echoes. Blind phases can



(A) TYPICAL PHASE DETECTOR CHARACTERISTIC

(B) PULSE SHAPES AT PHASE DETECTOR OUTPUT WITH COHO DETUNED $\frac{1}{2\tau}$

FIG. 10. *Phase detection.*

be eliminated by detuning the coho frequency, so that the resultant beats between signal and reference oscillators drive the detector through all possible phases during each pulse (Fig. 10b). For a 1 μs pulse, a minimum detuning of $\frac{1}{2}$ Mc/s would be necessary.

The spectrum of the phase detector output is shifted by this amount so that the video filter following should have a pass-band which is offset from zero frequency for optimum signal to noise ratio. The coherent low i.f. (c.l.i.f.) signal is then passed to the canceller, which should have a video bandwidth of $1.5/\tau$ to $2/\tau$ to accommodate the wider signal spectrum. The excess bandwidth requirement prevents c.l.i.f. operation in some short-pulse radars, but its use is recommended where practicable.

3.5. *Frequency Modulation*

In certain m.t.i. equipments, the delay line carriers are frequency modulated.[20] The outstanding advantage of this method is that the modulated carrier can be limited after amplification so that the demodulated output is independent of changes in the delayed carrier level. If the undelayed carrier is at a different frequency, subtraction of the delayed and undelayed signal information can take place in a frequency mixer, the resultant being

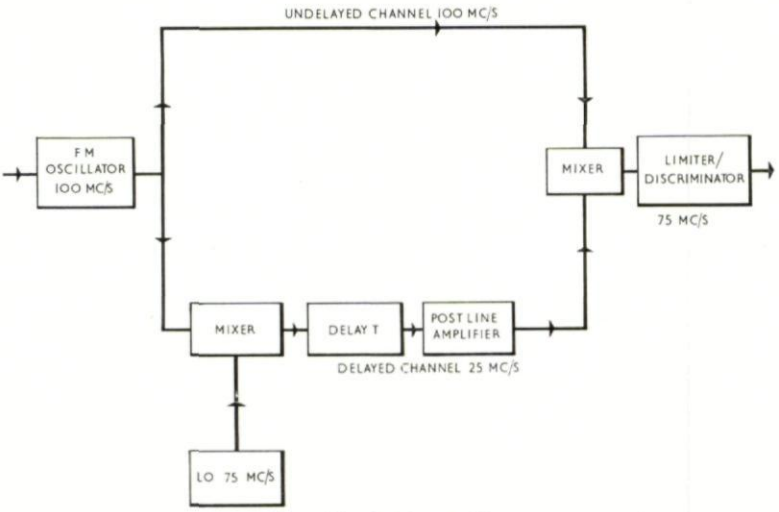


FIG. 11. *F.M. canceller.*

free from errors due to amplitude distortion (Fig. 11). F.M. cancellers have proved to be extremely reliable in service, but present certain special problems at the design stage.

The frequencies used have to be chosen with extreme care to prevent spurious intermodulation products. Carrier-to-noise ratio is not normally in question so that low-index modulation can be used but adequate bandwidth (say $4/\tau$) should be available to prevent transient distortion at the pulse edges. In a cascaded canceller it has been preferred from both design and field maintenance viewpoints to demodulate after each stage of cancellation and transfer the signal as bipolar video (Fig. 12).

Feedback video can also be added in, between first and second cancellation circuits if required.

In delay circuits for staggered p.r.f., f.m. carriers are used with the same advantages as in the canceller; intermodulation products are seldom of importance, as delayed and undelayed channels are at the same frequency. Again it is recommended to demodulate before feeding the signals into the canceller (Fig. 13).

3.6. *M.T.I. from a Moving Platform*

It is often required to cancel clutter in situations where there is relative motion between the radar platform and the clutter. A typical case is a shipborne search radar looking over the land.

Compensation for the ships' motion, by adding a correcting frequency to the coho output (Fig. 14) is the conventional solution[21] and has been described prior to 1948. The difficulty of this method lies in deriving an accurate control signal for the shift frequency, rather than in the frequency correction circuits themselves. The control signal must be computed from aerial bearing, ships true motion and aerial velocity due to ships roll and pitch.

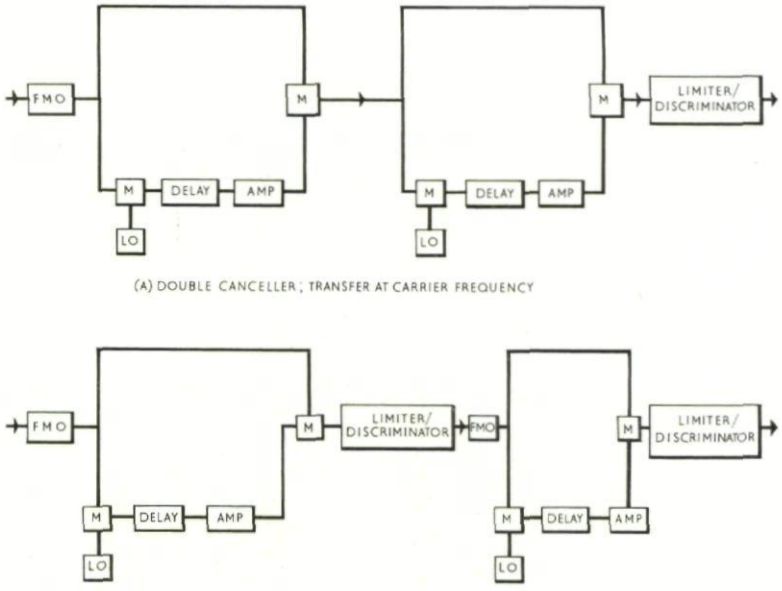


FIG. 12. *Cascaded f.m. cancellers.*

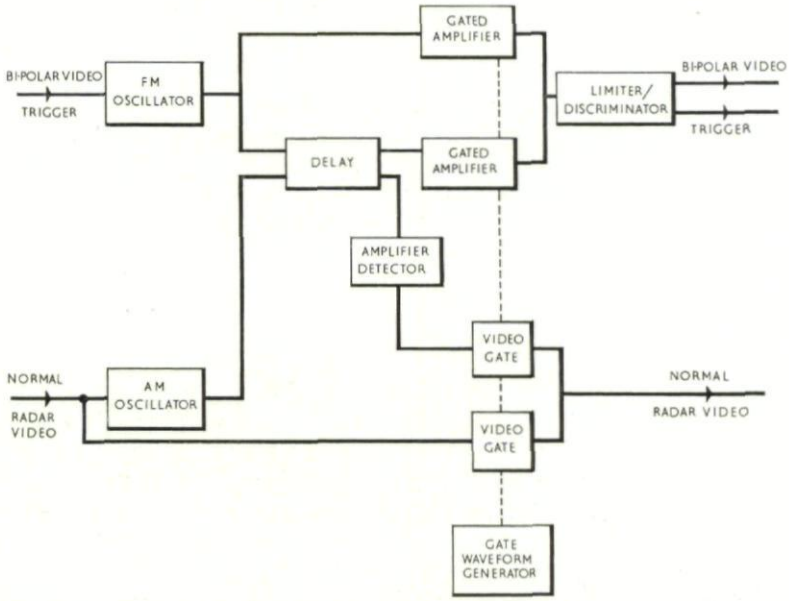


FIG. 13. *Practical staggered p.r.f. system.*

Non-coherent m.t.i. has no sensitivity to targets in the clear, and thus must be selected by a clutter operated gate. The gate circuits generally allow an annoyingly large number of clutter spots to appear on the display.

Two further techniques will be mentioned here, which may form the basis of more effective systems.

In the first, the i.f. clutter return is applied to the coherent oscillator which is then pulled into phase with the clutter. Delay in the locking channel

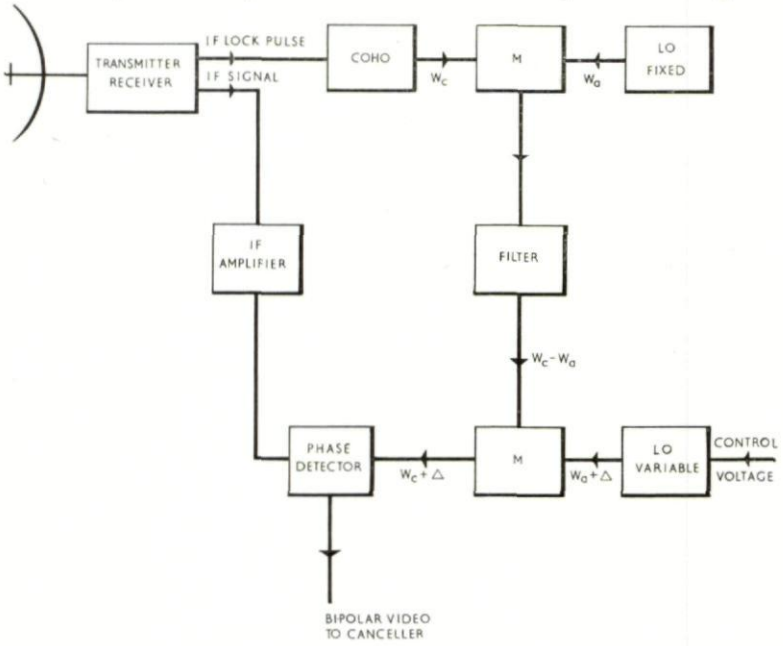


FIG. 14. *Velocity compensated coho.*

prevents a stationary phase condition from being established for point moving targets (Fig. 15).

The second "clutter-locked" system is associated with an i.f. canceller (Fig. 16).

The coho frequency w_c is mixed with the output of a voltage tuned oscillator at the delay line carrier frequency w_d. The difference frequency $w_c - w_d$ is subtracted from the limited i.f. signal w_s giving $w_s - w_c + w_d$, a coherent signal train which is fed to delayed and undelayed channels of the canceller, which are than combined in a phase detector. By adjustment of w_d the phase detector output can be brought to zero for stationary clutter.

Moving objects will then give a positive or negative output from the phase detector dependent on the magnitude and sense of their velocity, and independent of their initial phase. This latter property is an important point of difference from the conventional delay line canceller.

With an i.f. canceller, the clutter residue can be integrated and used to control w_d, so that the phase detector output is held at zero by degenerative

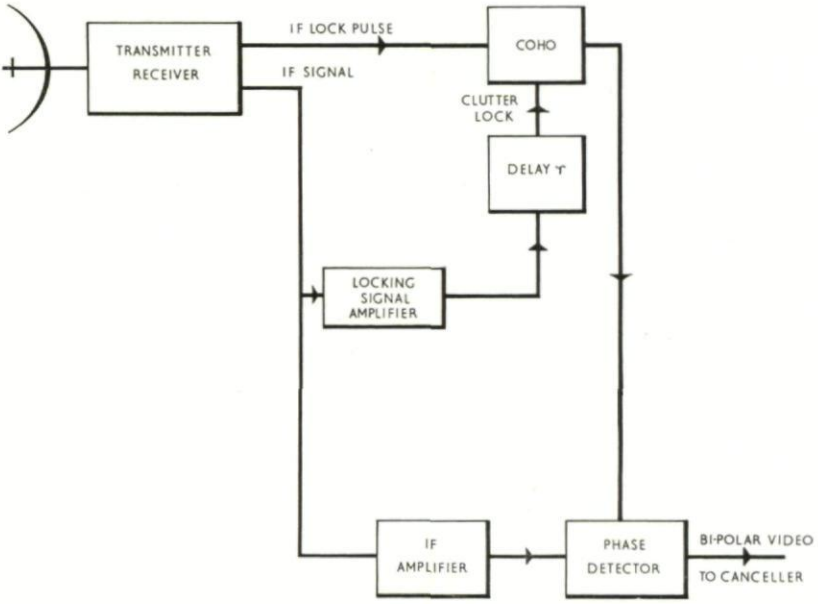


FIG. 15. *Clutter locked coho.*

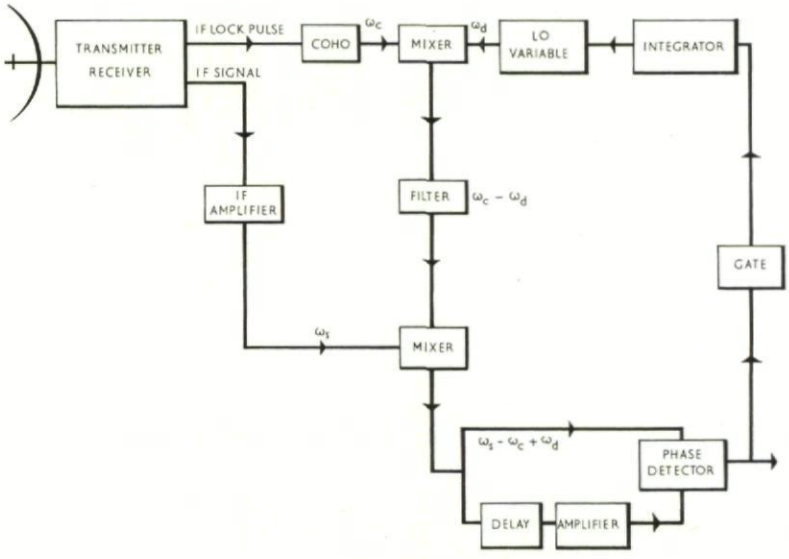


FIG. 16. *I.F. canceller with clutter lock.*

feedback. Such a feedback system has the valuable property that it will also cancel moving clutter, w_d taking up a different value to maintain the zero phase condition.

4. M.T.I. MEASUREMENTS

The usual measure of overall performance is the sub-clutter visibility (s.c.v.). This is defined as the ratio of fixed target power to the power of a moving target (at optimum speed) superimposed on the fixed target and just visible in fixed target residue.

The moving target is usually considered "just visible" when it reaches twice the amplitude of clutter residue as observed on an A-scope display.

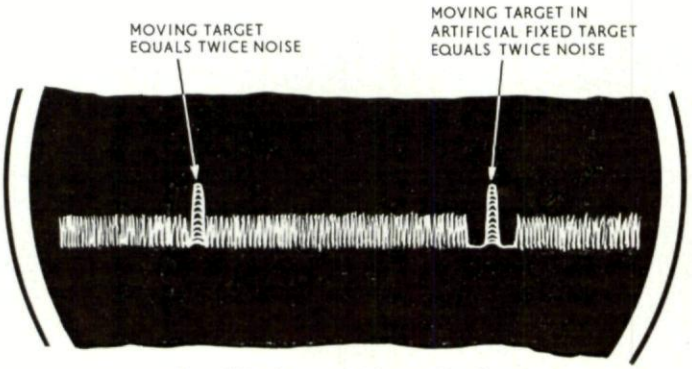


FIG. 17. *A-scope with test signals.*

The operational s.c.v., that is with real clutter echoes and a rotating aerial, can be measured directly only by a series of flight trials, regions of good and bad target visibility being plotted and compared with known strengths of aircraft and clutter echoes. This is a lengthly and expensive procedure.

Fortunately, reliable s.c.v. measurements can be made by using the system noise as a "transfer standard." The gain of the m.t.i. receiver is normally adjusted that the noise appearing on the display equals in brightness the clutter residue. All clutter reaching the i.f. limit level will paint at this brightness, independent of its amplitude at the receiver input.

If a strong artificial fixed target is injected into the system (at r.f. or i.f.) without disturbing the general noise level, the i.f. limiting action will cause the system noise to be suppressed at the range of the fixed target (Fig. 17). An optimum speed artificial moving target is now superimposed on the fixed target and is adjusted in amplitude so that its response (as monitored on an A-scope) rises up out of the hole made in the noise by the fixed target, and peaks to twice the shoulder noise level. The ratio (in db) between fixed and moving targets is noted. The assumption is made that if the artificial fixed targets were as poorly cancelled as are the real fixed targets, the residue would reach the observed system noise level, because the real fixed target residue reaches that level on the p.p.i. The moving target, at the same relative level which was noted, would then be just visible in the fixed target residue.

It follows that the fixed to moving target ratio determined is equal to the sub-clutter visibility.

One measuring procedure[22] called for an echo box to simulate the fixed target, and a pulsed r.f. signal generator to simulate the moving target.

The m.t.i. evaluator[23] provided pulsed fixed and moving targets at intermediate frequency. More recent m.t.i. test sets provide simulated targets at radio frequency which can be injected into the aerial feeder of an operating radar.

In the absence of special equipment, the overall m.t.i. performance may be judged by measuring the pre-cancellation limit level, that is the ratio of peak signal to noise at the input of the canceller. This ratio is approximately equal to the cancellation ratio (c.r.), the amount by which fixed targets are reduced in the process of cancellation. The c.r. is commonly taken to be 6 db greater than s.c.v.

Some m.t.i. systems have a built-in reflection delay line which produces a train of coherent i.f. pulses by delaying the coho lock pulse. These pulses are fed to the canceller in the test condition, and correct cancellation indicates satisfactory transmitter, receiver and canceller performance. The stalo is not tested by this method, and special test sets are now marketed for measuring the stalo frequency deviation. These are essentially superheterodyne f.m. receivers with stable crystal controlled local oscillators. Frequency deviations of a few parts in 10^9 are clearly indicated, and the disturbing waveform monitored to indicate the cause of instability.

Time jitter relative to a system pre-pulse can be adequately measured using modern oscilloscopes but absolute p.r.f. stability over a millisecond or more can best be determined with respect to a stable delay line as a standard, or by using a pulse jitter tester.[14]

The latter instrument contain a c.w. oscillator and display deviations of the pulse from the zero-crossings of the oscillator output either on an oscilloscope (as in the m.t.i. evaluator) or a direct-reading meter.

The various forms of modulated reflector are an important class of m.t.i. test equipment. These are passive moving targets, deployed in the field so as to give spot echoes similar to those from corner reflectors on non-m.t.i. displays. The devices include mechanically rotated corner reflectors, and modulated ferrites or diodes at the focus of microwave reflectors. The latter, driven by semi-conductor circuits, are capable of long life in the field without attention. They give a ready check to the operator of equipment sensitivity and range and bearing accuracy of the radar and can serve as runway or similar approach markers.

5. THE STATE OF THE ART IN COHO-STALO M.T.I.

Equipments offered are mainly in the *L* and *S*-bands and have sub-clutter visibilities between say, 15 and 30 db. At the higher frequencies (3000 Mc/s) and lower p.r.f. (250 p.p.s.) the cancellation is limited by clutter fluctuation alone to 20–30 db (depending on conditions).

The canceller itself is capable of 36–40 db, and by appropriate choice of velocity characteristics the effect of scanning should be no more than 30 db for pulses per beamwidth greater than 10.

Stalo design still presents many engineering problems if the exacting specifications of the present day are to be met. Under field service conditions the stalo may be the limiting factor in performance but the designer has failed if this is the case.

The magnetron is the device which we must either improve or discard in developing the next generation of m.t.i. radars. As the design of other components is improved the effect of residual f.m., usually random in distribution, is left as the principal factor determining m.t.i. performance.

6. AIRBORNE M.T.I.

The applications of airborne m.t.i. (a.m.t.i.) can be divided into two groups, air-to-ground radar and air-to-air or airborne early warning (a.e.w.) radar. Detection of ground targets requires medium range and very good

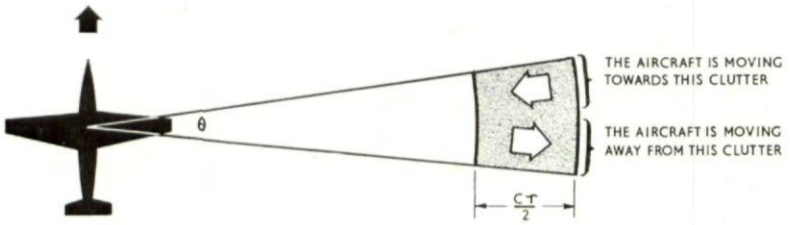


FIG. 18. *Clutter illuminated by a.m.t.i.*

resolution, hence the use of frequencies in *X*-band and above. Some clutter return is always present so that non-coherent m.t.i. can be used. The equipment is simplified in that no stalo is required, but the transmitter still requires good time and frequency stability. The clutter spectrum in a.m.t.i. is broadened because the patch of clutter illuminated by the pulse (Fig. 18) consists of elements having different velocities relative to the aircraft. This effect, due to the linear motion of the aircraft, is greater than the effect of aerial rotation, especially in directions perpendicular to the aircraft's track. The techniques of double cancellation and velocity shaping have accordingly been applied to a.m.t.i.

Airborne ground-mapping radars do not use m.t.i.; their function is that of fixed target indication, but m.t.i. principles have been used to improve the resolution of sideways-looking radars.[22] The radar beam is directed broadside from the aircraft and the received coherent video signals are fed into a delay line integrator. Such an integrator[26] has velocity characteristic which is the inverse of an m.t.i. canceller, the response being sharply peaked at zero frequency and harmonics of the p.r.f. The integrator thus rejects signals received from targets slightly off the axis of the beam, as these have a small

resolved component of velocity either towards or away from the aircraft, and enhances returns from targets "on axis."

The result is a considerable narrowing of the effective aerial beamwidth, or an increase in the aerial horizontal aperture, hence the name "synthetic-aperture radar." The effective aperture is in fact equal to the distance travelled by the aircraft in the storage time of the integrator.

Airborne early warning calls for lower radio frequencies than the *X*-band and coverage beyond the horizon, so that a clutter background is not always present.

Coherent m.t.i. is thus indicated, and the most practicable forms of compensation for platform motion are the clutter-lacked systems of section 3.6.

7. MORE COMPLEX TRANSMITTED WAVEFORMS

Within the last ten years, high power r.f. amplifiers have been used to an increasing extent in the output stages of radar transmitters. Klystrons, grid controlled tubes (at the lower frequencies), and more recently crossed-field amplifiers and travelling wave tubes have been developed, giving a new degree of freedom in the choice of transmitted signal. Improved stability is also possible by deriving the transmitted frequency from a stable master oscillator.

7.1. *Twin-pulse m.t.i.*

High power, long range radar needs a low p.r.f., and hence m.t.i. performance is limited by clutter fluctuation, the clutter becoming to some extent de-correlated over a comparison interval of several milliseconds. In twin-pulse m.t.i., pairs of pulses, separated by a few hundred microseconds, are radiated, and phase comparison takes place between pulses of a pair the group p.r.f. being unimportant.

Clutter fluctuation, scanning effect and velocity response are as for a radar with p.r.f. equal to the inverse of the pulse spacing. One restriction on the design is that the pulse spacing must not be less than the duration of the longest clutter echoes to be handled, e.g. a 600 μs pulses pacing could deal with up to 48 miles of clutter.

The i.f. canceller of Section 3.6, is particularly suited to twin-pulse m.t.i.

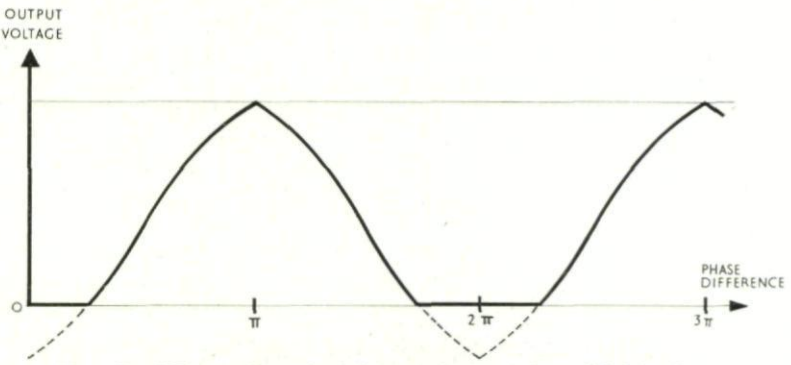


FIG. 19. *I.F. canceller: phase detector characteristic modified by clipping.*

In its simplest form the velocity characteristic is that of a single canceller, which may not give enough clutter rejection in a high data rate equipment. However, the velocity response could be modified by operating on the phase detector characteristic; for instance clipping the negative going peak of the detector output (Fig. 19) will result in a widening of the blind speed gap in the velocity response, and thus improved cancellation.

7.2. *Two Frequency m.t.i.*

The principle of two frequency m.t.i. has been proposed a number of times in the last few years without finding a practical application. The target is illuminated simultaneously at two radio frequencies and two coherent receivers are interconnected so as to produce a video output dependent on the phase difference between the received signals at the two frequencies. The velocity response of the system is similar to that of an m.t.i. radar transmitting a frequency which is the difference between the two radio frequencies. If the latter are 2800 and 3000 Mc/s and the p.r.f. 600 p.p.s. the first blind speed would be approximately 750 knots. The effect of scanning and of internal fluctuation of clutter is expected to be the same as for the individual radio frequencies.

7.3. *M.T.I. and Pulse Compression*

The requirements of pulse compression and m.t.i. are in principle compatible. In one form of pulse compression radar (Fig. 20) the transmitted waveform is generated by passing a short intermediate frequency pulse

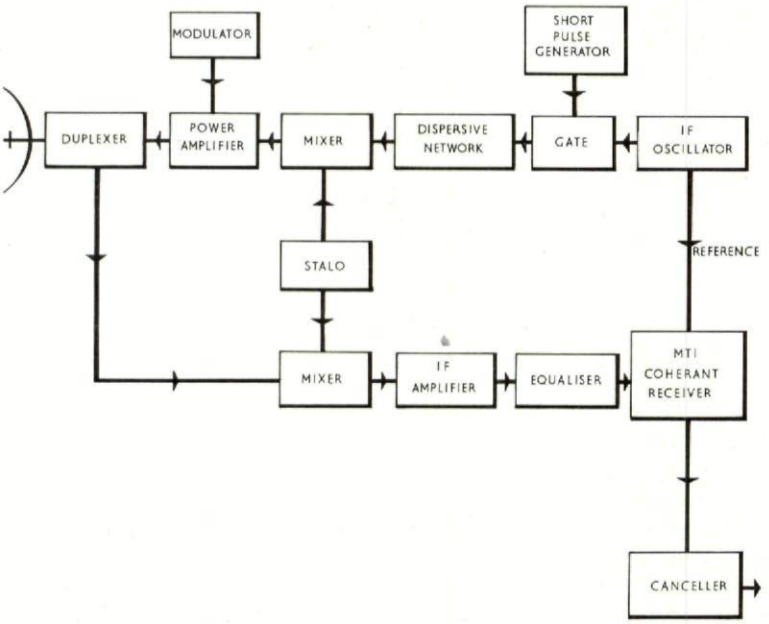


FIG. 20. *Pulse compression + m.t.i.*

through a dispersive network. The resulting long pulse is converted to r.f., amplified and transmitted. The received target echoes are passed, at i.f., through an equalising network which "compresses" the pulses again. If all the signal handling circuits are linear the output of the equaliser is similar to the received signal from a radar transmitting the original short pulse. In particular, the pulse-to-pulse shift of the compressed signals received from fixed targets will be zero, provided the transmitted frequencies are derived from a stable source.

The dynamic range of an m.t.i. canceller does not usually exceed 40 db, due to noise and spurious signals in the delay line circuits. The signal to noise ratio at the input of the radar receiver may be 80–100 db, so that some limiting is inevitable. If this takes place before pulse compression there will be intermodulation between overlapping long pulses of clutter, and resolution will be lost. For optimum detection of moving point targets in the presence of clutter the limiter, phase detector and canceller of the m.t.i. should follow the pulse compression circuits. Pulse compression will improve the target to clutter ratio as compared with a long pulse transmission at constant frequency, and the m.t.i. will then reject fixed echoes, giving a twofold discrimination against stationary clutter.

If limiting takes place before pulse compression, the system has a "constant false alarm" property which may be useful in special circumstances. The method of generating the transmitted waveform is not restricted to that discussed above but if the source is a frequency swept oscillator it may be impossible to meet the stability requirements for m.t.i.

The m.t.i. circuits must be able to handle the bandwidth of the compressed pulse; approximately 10 Mc/s is believed to be the maximum that can be achieved in the present state of m.t.i. development.

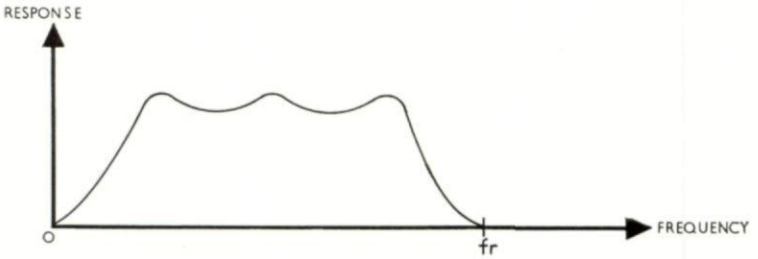
The discussion above also applies generally to systems of pulse compression in which a coded sequence of discrete frequencies and phases is transmitted. Equalisation in the receiver involves selective delay circuits in which elements of the code are superimposed, the output being a pulse equal in length to a single element.

8. PULSE-DOPPLER RADAR

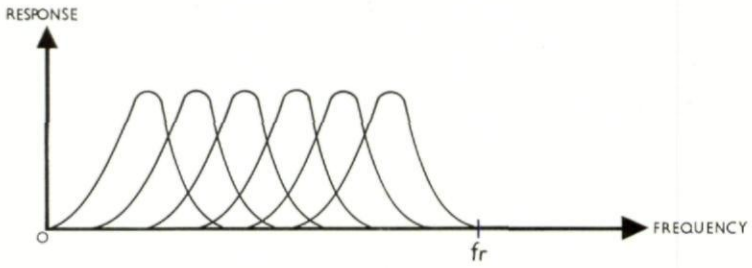
If the output of a coherent radar receiver is passed through a video bandpass filter which rejects adjacent harmonics of the p.r.f. but accepts frequencies lying between these harmonics (Fig. 21a) then stationary clutter will be rejected and moving targets accepted. Because the filter has a relatively narrow bandwidth compared with the pulse spectrum width, range resolution is lost. Signal to noise ratio is also lost, because noise in adjacent range elements is added (collapsing loss).

For this reason, the range resolution element must be defined by a video gate before the filter.

A number of m.t.i. radars based on the pulse doppler principle have been developed,[27] mainly for military applications Possibly the simplest is a tracking radar in which a single range gate follows the target echo (Fig. 22.)



(A) CLUTTER REJECTION FILTER



(B) FILTER BANK

FIG. 21. *Pulse Doppler filter responses.*
(*a*) *Clutter rejection filter.*
(*b*) *Filter bank.*

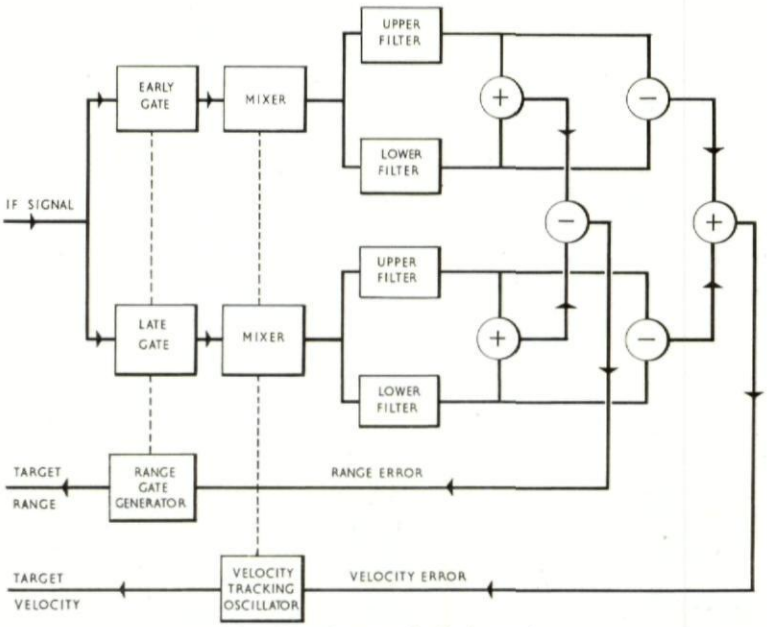


FIG. 22. *Range and velocity tracker*

The gate output is fed to a narrow band filter which is automatically tuned to track the target velocity. In the most elaborate multiple-target systems an array of range gates is used, each gate feeding a separate bank of Doppler filters (Fig. 23). Division of the Doppler band between a number of filters reduces the effective receiver bandwidth and improves the signal to noise ratio.

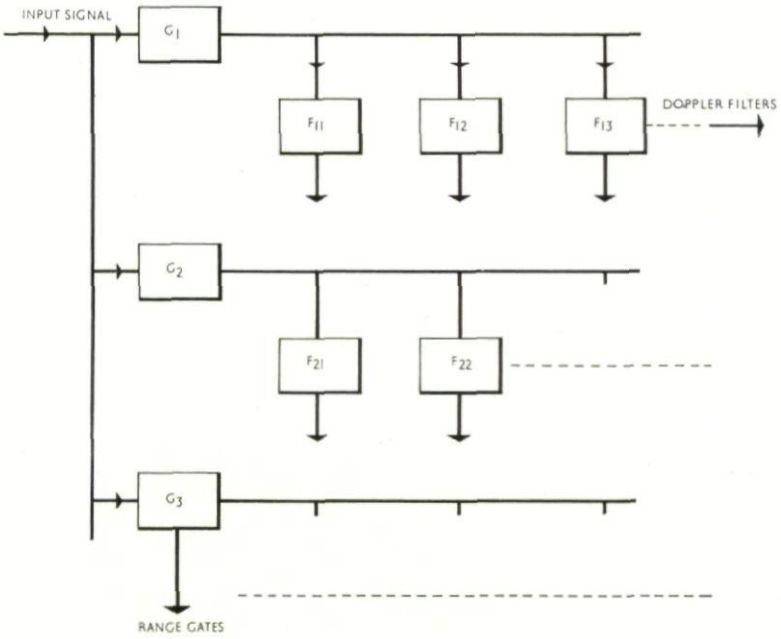


FIG. 23. *Array of range gates and filters.*

An interesting recent development is that of pulse Doppler signal processing by means of a time transformation technique, enabling a filter with rapid response to be used. The heart of the system is a recirculating video integrator with delay time $T + \Delta$ where T is the interpulse period. A train of n pulses, received during a target illumination time nT, appears at the integrator output with spacing Δ occupying a time $n\Delta$. The envelope of pulse amplitudes is compressed in time in the ratio T/Δ so that a Doppler frequency fd appears at the following filter as $fd \cdot T/\Delta$. As an example, suppose $n = 10$, $T = 1$ ms, $\Delta = 1$ μs. If $fd = 100$ c/s, one cycle will be received in time $nT = 10$ ms. After integration one Doppler cycle will occupy $n = 10$ μs the new frequency being 100 kc/s.

A frequency shift of this amount can be resolved within one range resolution element.

To summarise, pulse Doppler systems find their main application where the target illumination time (nT) is long. They are capable of higher clutter attenuation than other m.t.i. systems, but become more complex when

many range gates are installed for high resolution, each gate feeding a Doppler filter or bank of filters.

Because range and velocity are quantised, such systems are more suited to automatic detection than to visual displays. The maximum velocity which can be detected without ambiguity corresponds to a Doppler frequency of the p.r.f., while the choice of p.r.f. is itself restricted by range ambiguity.

9. AREA M.T.I.

The name "Area m.t.i." is given to systems which detect moving targets by their change in plan position from scan to scan. The principle has been accepted as desirable for a number of years; the response to velocities over an initial threshold is constant, and of all m.t.i.'s area m.t.i. makes the fewest demands on the performance of either radar or target.

However, due to limitations of technique, few area m.t.i. systems are in operational service.

9.1. *Storage Tubes*

In a storage tube m.t.i. the radar signals are stored from one scan to another in the form of electrostatic charges on a mosaic or mesh similar to that of a television camera tube. One of the earliest types of tube[28] used for m.t.i. was the "Radechon." A first requirement is for good resolution; the threshold velocity is the target speed at which the target moves from one resolution element to the next in one scan. The larger the size of resolution cell, the higher will be the threshold velocity although this can be reduced either by comparing over, say 4 or 5 scans or by presenting only part of the radar picture with an expanded scale on the storage tube.

Cancellation of permanent echoes can take place either outside or inside the tube; in the former case the stored information is read out by a separate electron beam and subtracted from the incoming radar signals. With internal cancellation the storage elements carry charges which are stabilised at level of the radar signals in one writing scan. The next scan will show a different charge pattern where moving targets are present, and the charge differences will be picked up either from the tube backing place or mask electrode, and after amplification will form the m.t.i. video signal.

9.2. *Display Storage Tubes*

A direct view storage tube has recently been demonstrated in which, by change of electrode potential, a previously written charge pattern could be erased by further "negative" writing. It is felt that this device has an application in the future for permanent echo suppression.

A storage tube in a different sense was the display tube proposed some years ago with a two colour phosphor, the second colour being excited only after repeated electron bombardment, and thus indicating fixed echoes.

9.3. *Digital Storage*

Some radar data processing equipment records target plots in a ferrite core matrix or similar computer store. The movement of targets from one

address to another can readily be detected by the computer, and permanently filled plot positions ignored during the processing. This system is vulnerable, however, in an environment in which many false reports are generated, thus filling the storage capacity of the computer.

10. A LOOK AT THE FUTURE

In the immediate future it appears of prime importance to bring reliability and easy maintenance to conventional m.t.i. systems; the use of semiconductors in new production equipment will go a long way towards realising these aims. Looking further ahead, it is expected that i.f. cancellation systems will be seen more frequently in "conventional" m.t.i. and also in twin-pulse systems with pulse compression. The ultimate development from twin-pulse m.t.i. will be moving target detection within a single pulse.

Another promising method is to quantise the coherent video signals into sufficient amplitude levels (100 for 40 db cancellation) and to code these into binary words, subtracting in a computer. First experiments on these lines have already taken place.

Recent improvements in the resolution of storage tubes will make possible new standards of performance in area m.t.i. equipment. Finally, in systems of extremely fine resolution, the target movement between pulses should be detectable, permitting direct video cancellation. Sufficient bandwidth would be achieved using sampling techniques, with parallel handling of binary coded signals in relatively cheap wire delay lines.

11. ACKNOWLEDGEMENT

Acknowledgements are due to the Management of Decca Radar Limited for permission to publish this paper, and to Mr. G. Oulsnam of Decca Radar Limited for programming the computer for the work described in the appendix.

APPENDIX

An Analysis of Two- and Three-Period Staggered P.R.F. M.T.I. Systems

1. *General*

Perlman[19] describes a two-period stagger applied to a single cancellation m.t.i. system. For a main delay line of length T and a stagger delay of ρ the system response is the average of the responses to separate pulse trains of period $T + \delta$ and $T - \delta$.

In a double cancellation system pulse pairs separated by $T + \delta$ and $T - \delta$ are subtracted in the first canceller and the differences fed into the second canceller so that the system responses to the two pulse separations cannot be considered separately.

In a three-period stagger system, delay lines T, δ are switched to give periods $T + \delta$, T, and $T - \rho$, so that a greater diversity of pulse spacing is obtained with little increase in complexity of the equipment.

2. *Uniform P.R.F.*

The Doppler frequency associated with a target speed V and radar wavelength λ is $2V/\lambda$. Let the output of a coherent detector consist of a train of pulses modulated with the sinusoidal envelope $E = \sin 2\pi f\, dt$.

If uniform (non-staggered) p.r.f. is considered, the amplitudes of successive pulses are derived by sampling the envelope at times $t_n = T_0 + nT$, and the amplitude of the nth sample is $A_n = \sin 2\pi fd(t_0 + nT)$.

3. *Two-period Stagger*

The pulse envelope at the input to the stagger circuit is as before, but the envelope is sampled at times $t_0 + nT$ for n even, and $T_0 + nT + \delta$ for n odd, i.e. $t_n = t_0 + \frac{\delta}{2} + nT + (-1)^n \frac{\delta}{2}$ for all n.

The pulse train is realigned to uniform spacing before cancellation, even pulses being delayed by a time δ. The amplitude of the nth pulse is still

$$A_n = \sin 2\pi fd\left(t_0 + \frac{\delta}{2} + nT + (-1)^n \frac{\delta}{2}\right)$$

$$= \sin 2\pi fd\left(t_1 + nT + (-1)^n \frac{\delta}{2}\right)$$

putting $t_1 = t_0 + \frac{\delta}{2}$

$$\text{Expanding, } A_n = \sin 2\pi fd(t_1 + nT) \cdot \cos 2\pi fd(-1)^n \frac{\delta}{2}$$

$$+ \cos 2\pi fd(t_1 + nT) \cdot \sin 2\pi fd(-1)^n \frac{\delta}{2}$$

Now $\sin(-\theta) = -\sin\theta$, and $\cos(-\theta) = \cos\theta$

$$\text{so } A_n = \sin 2\pi fd(t_1 + nT) \cdot \cos 2\pi fd \frac{\delta}{2} + (-1)^n \cos 2\pi fd(t_1 + nT) \sin 2\pi fd \frac{\delta}{2}$$

for n integral, $(-1)^n = \cos n\pi$.

$$A_n = \sin 2\pi fd(t_1 + nT) \cos 2\pi fd \frac{\delta}{2} + \cos 2\pi fd(t_1 + nT) \cos \pi n.\ \sin 2\pi fd \frac{\delta}{2}$$

Express the product of cosines as the sum of two cosines:

$$A_n = \sin 2\pi fd(t_1 + nT) \cos 2\pi fd \frac{\delta}{2}$$

$$+ \frac{1}{2} \cos 2\pi fd \left[t_1 + n\left(T + \frac{1}{2fd}\right)\right] \sin 2\pi fd \frac{\delta}{2}$$

$$+ \frac{1}{2} \cos 2\pi fd \left[t_1 + n\left(T - \frac{1}{2fd}\right)\right] \sin 2\pi fd \frac{\delta}{2}$$

$$\text{Put } \cos 2\pi f_1(t_1' + nT) = \cos 2\pi fd \left[t_1 + n\left(T + \frac{1}{2fd}\right)\right]$$

$$\text{Thus } f\, dt_1 = f_1 t_1'$$

and $$f_1 = fd\left(1 + \frac{2}{fdT}\right) = fd + \frac{1}{2}fr \qquad \left(fr = \frac{1}{T}\right)$$

Similarly, $$f_2 = fd\left(1 - \frac{1}{2fdT}\right) = fd - \frac{1}{2}fr$$

Finally, $$A_n = \sin 2\pi fd(t_1 + nT) \cdot \cos \pi f\, d\delta + \cos 2\pi\left(fd + \frac{1}{2}fr\right)(t_1' + nT) \cdot \frac{1}{2}\sin \pi fd\delta + \cos 2\pi\left(fd - \frac{1}{2}fr\right)(t_1' + nT) \cdot \frac{1}{2}\sin \pi fd\delta.$$

The pulse envelope at the input to the canceller has been resolved into three components:—

(1) A sinusoidal envelope of frequency fd, amplitude $\cos \pi fd\delta$

(2) A cosinusoidal envelope of frequency $fd + \frac{1}{2}fr$, amplitude $\frac{1}{2}\sin \pi fd\delta$

(3) A similar envelope of frequency $fd - \frac{1}{2}fr$.

The frequency response $G(fd)$ of the canceller is periodic in fr, so that $G(fd + \frac{1}{2}fr) = G(fd - \frac{1}{2}fr)$, and the two "sideband" responses can be added together.

The r.m.s. output of the canceller is thus:—

$$V_{\text{out}} = [G^2(fd) \cdot \cos^2 \pi f\, d\delta + G^2(fd + \tfrac{1}{2}fr) \sin^2 \pi f\, d\delta]^{1/2}$$

Note that when fd corresponds to a blind speed, $(fd + \frac{1}{2}fr)$ corresponds to an optimum speed so that the second term fills in the minima of the first term.

For double cancellation, $G(fd) = \sin^2 \pi f\, dT$ and r.m.s. cancelled output =

$$(\sin^4 \pi f\, dT \cos^2 \pi f\, d\delta + \cos^4 \pi f\, dT \sin^2 \pi f\, d\delta)^{1/2} \qquad (1)$$

4. *Three-period Stagger*

Here the nth sample is taken at $t_0 + nT + \delta$,

the $(n+1)th$ at $t_0 + (n+1)T$,
$(n+2)th$ at $t_0 + (n+2)T$,
$(n+3)th$ at $t_0 + (n+3)T + \delta \cdots$

giving successive periods $T - \delta$, T, $T + \delta$. A function is required having values $1, o, o, 1 \cdots$ for integral values of the variable n. The function

$$\left(\frac{1}{3} + \frac{2}{3}\cos\frac{2n\pi}{3}\right)$$

has this property.

Now $\left(\frac{1}{3} + \frac{2}{3}\cos\frac{2n\pi}{3}\right) = \frac{1}{2}\left[I + \left(\frac{4}{3}\cos\frac{2n\pi}{3} - \frac{1}{3}\right)\right]$ where $\left(\frac{4}{3}\cos\frac{2n\pi}{3} - \frac{1}{3}\right)$

takes the values $+1, -1 \cdot -1, +1, \cdots$ for n integral.

Let the amplitude of the nth pulse be:—

$$A_n = \sin 2\pi fd\left\{t_0 + nT + \frac{\delta}{2}\left[1 + \left(\frac{4}{3}\cos\frac{2n\pi}{3} - \frac{1}{3}\right)\right]\right\}$$

$$\text{make } t_1 = t_0 + \frac{\delta}{2}$$

$$A_n = \sin 2\pi fd\left\{t_1 + nT + \frac{1}{2}\left(\frac{4}{3}\cos\frac{2n\pi}{3} - \frac{1}{3}\right)\delta\right\}$$

$$= \sin 2\pi fd(t_1 + nT)\cos\pi fd\left(\frac{4}{3}\cos\frac{2n\pi}{3} - \frac{1}{3}\right)\delta$$

$$+ \cos 2\pi fd(t_1 + nT)\sin\pi fd\left(\frac{4}{3}\cos\frac{2n\pi}{3} - \frac{1}{3}\right)\delta$$

Since $\sin(-\theta) = -\sin\theta$ and $\cos(-\theta) = \cos\theta$

$$A_n = \sin 2\pi fd(t_1 + nT)\cos\pi f\,d\delta$$

$$+ \cos 2\pi fd(t_1 + nT)\left(\frac{4}{3}\cos\frac{2n\pi}{3} - \frac{1}{3}\right)\sin\pi f\,d\delta$$

$$= \sin 2\pi fd(t_1 + nT)\cos\pi f\,d\delta - \frac{1}{3}\cos 2\pi fd(t_1 + nT)\sin\pi f\,d\delta$$

$$+ \frac{4}{3}\cos 2\pi fd(t_1 + nT)\cos\frac{2n\pi}{3}\sin\pi f\,d\delta.$$

Expanding the product of cosines, and putting $fr = \frac{1}{T}$ we have finally:

$$A_n = \sin 2\pi fd(t_1 + nT)\cos\pi f\,d\delta - \frac{1}{3}\cos 2\pi fd(t_1 + nT)\sin\pi f\,d\delta$$

$$+ \frac{2}{3}\cos 2\pi\left(fd + \frac{fr}{3}\right)(t_1 + nT)\sin\pi f\,d\delta$$

$$+ \frac{2}{3}\cos 2\pi\left(fd - \frac{fr}{3}\right)(t_1 + nT)\sin\pi f\,d\delta$$

Note in this case that the "sidebands" are spaced $\pm\frac{1}{3}$ of the p.r.f. from the Doppler frequency; the canceller has a different response to each sideband.

The r.m.s. canceller output for three-period stagger is

$$V_{\text{out}} = \left[G^2(fd)\left(\cos^2\pi f\,d\delta + \frac{1}{9}\sin^2\pi f\,d\delta\right) + G^2\left(fd + \frac{fr}{3}\right)\frac{4}{9}\sin^2\pi f\,d\delta\right.$$

$$\left. + G^2\left(fd - \frac{fr}{3}\right)\frac{4}{9}\sin^2\pi f\,d\delta\right]^{1/2}$$

For double cancellation, $G(fd) = \sin^2\pi f\,dT$ and the r.m.s. canceller output is:

$$\left[\sin^4\pi f\,dT\left(\cos^2\pi f\,d\delta + \frac{1}{9}\sin^2\pi f\,d\delta\right) + \sin^4\left(\pi f\,dT + \frac{\pi}{3}\right)\frac{4}{9}\sin^2\pi f\,d\delta\right.$$

$$\left. + \sin^4\left(\pi f\,dT - \frac{\pi}{3}\right)\frac{4}{9}\sin^2\pi f\,d\delta\right]^{1/2} \qquad (2)$$

TABLE I

M.T.I. VELOCITY RESPONSES

Tables of the three-period stagger ratio function

$$V^{out} = \{\sin^4 \pi x - \tfrac{4}{3} \sin^2 \pi ax \,.\, \sin^4 \pi x + \tfrac{1}{2} \sin^2 \pi ax\}^{1/2}$$

Where $a = M/N$

$M = 1 \quad N = 8$

$F(X)$	$20_{\log_2}(X)$	X
0.099	−14.04	0.100
0.349	−3.13	
0.654	2.33	
0.896	5.07	
0.984	5.88	
0.887	4.97	0.600
0.650	2.29	
0.390	−2.16	
0.260	−5.68	0.900
0.256	−5.81	0.925
0.271	−5.33	1.000
0.308	−4.22	
0.435	−1.20	
0.641	2.16	
0.810	4.20	
0.862	4.73	1.500
0.784	3.90	
0.633	2.05	
0.513	0.22	
0.483	−0.30	1.888
0.484	−0.29	1.900
0.500	0.00	
0.522	0.37	
0.563	1.02	
0.620	1.87	
0.657	2.37	2.400
0.651	2.29	
0.623	1.92	
0.612	1.75	2.690
0.612	1.75	2.700
0.624	1.93	2.800
0.641	2.16	
0.653	2.32	
0.662	2.44	
0.654	2.33	
0.602	1.61	3.300
0.509	0.16	
0.445	−1.00	3.500
0.445	−1.02	3.511
0.492	−0.15	
0.598	1.56	3.700
0.677	2.63	
0.704	2.98	
0.707	3.01	4.000

The expressions (1) and (2) represent the response of the m.t.i. system to targets of different speeds. The expressions can be computed and plotted as functions of the angle $\pi f_d T$ for given values of δ/T so that a set of curves are obtained which can be scaled to give the velocity characteristics of any radar of given wavelength and p.r.f.

It is convenient to define

$$x = f_d T ;$$

and

$$a = \frac{\delta}{T} = \frac{m}{n}$$

so that for two period stagger with double cancellation the r.n.s. cancelled output is

$$V_{\text{out}} = (\sin^4 \pi x \cos^2 \pi a x + \cos^4 \pi x \sin^2 \pi a x)^{1/2} \tag{3}$$

For three period stagger the corresponding output is

$$\begin{aligned} V_{\text{out}} = [&\sin^4 \pi x(\cos^2 \pi a x + \tfrac{1}{9} \sin^2 \pi a x) \\ &+ \tfrac{4}{9} \sin^4 \pi(x + \tfrac{1}{3}) \sin^2 \pi a x \\ &+ \tfrac{4}{9} \sin^4 \pi(x - \tfrac{1}{3}) \sin^2 \pi a x]^{1/2} \end{aligned}$$

For the purposes of computation this latter expression is more conveniently written as

$$V_{\text{out}} = (\sin^4 \pi x - \tfrac{4}{3} \sin^2 \pi a x \sin^4 \pi x + \tfrac{1}{2} \sin^2 \pi a x)^{1/2}$$

This function is shown plotted in Fig. 9b where the stagger ratio is 7 : 8 : 9, that is, with $a = 1/8$.

The function is also tabulated in the table for $x = 0.1\ (0.1)\ 4.0$ and also those values of x for which the function is a minimum. The second column of the Table gives the value of the function in dbs referred to 0.5 as 0 db.

The calculations for the above function were carried out on the Pegasus Mk. 11 Computer at Southampton University.

DISCUSSION

L. Botwin: I would like to comment on your last slide. We have been doing some basic work on range gated m.t.i. systems. Filters such as those of your last slide have been built on molecular chips which are completely integrated with series condensers and parallel resistors to ground. We have been able to use a planor transistor on the molecular chip to change the parallel resistance to ground by voltage control, and in this way vary the time constant by a factor of 100.

REFERENCES

1. Ridenour, L. N., *Radar System Engineering*, p. 667. MIT Series Vol. 1, McGraw-Hill 1947.
2. Bauer, P. A., Low temperature coefficient ultrasonic delay lines, *Solid State J*, **2,** 12, 23, 1961.

3. MAY, J. R. JR., Low loss 1000 μs ultrasonic delay lines, *Proc. Natl. Electronics Comp.* **11,** 786–790, 1955. Also *Trans IRE* **UE-7,** 1, 1960.
4. HUTSON, A. R., MCFEE, J. H., WHITE, D. L., *Phys. Rev.* **7,** 237 (1961).
5. Ref. 1 above, p. 659.
6. POUND, R. V., Frequency stabilisation of microwave oscillators, *Proc. IRE.* **35,** 1405 (1947).
7. TULLER, GALLOWAY and ZAFFERANCE, Recent developments in the frequency stabilisation of microwave oscillators, *Proc. IRE*, **36,** 794–800, 1948.
8. STEPHENSON, J. G., Designing stable triode microwave oscillators, *Electronics*, **184,** 1955.
9. MALLING, L. R., Phase stable oscillators, *Proc. IRE*, **50,** 7, 1656, 1962.
10. DAUKSHER, W. J., Stable local oscillators for *S*-band radar. *Electronics*, **29,** 179, 1956.
11. FORTINI, M. M., and VILMS, J., Solid state generator for microwave power, *Electronics*, **42,** 1959.
12. BENJAMINSON, A., Phase locking microwave oscillators to improve stability and FM. *Microwave Journal*, **6,** 1, 88, 1963.
13. REY, T. J. and PASCALAR, H. G., *Trans IRE*, **CS-10,** 2 202–8, 1962.
14. VAN VOORHIS, *Microwave Receivers*, Chap. 22, MIT Series Vol. 23, McGraw-Hill 1948.
15. ADLER, R., *Electronics*, 82, 1950.
16. Ref. 1 above, p. 644.
17. SOLOMON, K., A double delay and subtraction airborne canceller, *Proc. Conf. Military Electronics* (IRE), 235–240, 1958.
18. WHITE, W. O. and RUVOLN, A., Recent advances in the synthesis of comb filters, *IRE Nat. Conv. Record*, **5,** pt. 2, 186–199, 1957.
19. PERTMAN, S. E., Stagger rep(etition) rate fills radar-blind spots, *Electronics*, **31,** 47 82–85, 1958.
20. MCKEE, D. A., An FM. M.T.I. cancellation system, *MIT Lincoln Lab. Tech. Report* 171, 1958.
21. See Ref. 1, p. 655.
22. F.A.A. Specification, FAA-R-1286.
23. WILKERSON, J. R., An M.T.I. evaluator, *IRE Nat. Conv. Record*, 1956.
24. FITCH, J. L. and BUSS, R. E., Measurement of time jitter in trains of video pulses, *Trans. IRE*, **PGI-3,** 23, 1954.
25. SHERWIN, C. W. *et al.* Some early developments in synthetic aperture radar systems, *Trans. IRE*, **MIL-6,** 2, 111–115, 1962.
26. MACFARLANE, A. G. J., Analysis of a type of comb filter, *Proc. Inst. Elect. Engrs.*, Paper 3121E, 1960; also **107A**, 39–52.
27. MINTZER, A. I., Advanced radar signal and data processing—111, *Space Aeronautics*, **39,** 4, 93–98, 1963.
28. JENSEN, A. S., The Radechon. A barrier grid storage tube, *RCA Rev.* **XVI**, 2, 179, 1955.

CHAPTER 10

LE RADAR SDS, RADAR DE SURVEILLANCE DESTINE A DETECTER LES OBJECTIFS MOBILES AU SOL: EXTENSION A LA DETECTION DES AVIONS VOLANT BAS

G. VAN DEN BROEK
Laboratoire Central de Télécommunications, Paris

An account is given of a radar designed to fulfill the following functions:

(a) *The detection of moving targets on the ground with a range of* 30 *Km for vehicles and* 15 *Km for a man.*
(b) *The measurement of bearing, elevation and distance of the target with a precision of* 5 *milliradians and* 40 *metres respectively.*
(c) *The identification of the nature of the targets (wheeled vehicles, tracked vehicles, pedestrians singly or in groups etc.).*
(d) *The identification of friendly targets.*

The radar operates in X-band with a 40 *Kw peak power, a p.r.f. of* 4 *kc/s and* 0.25 *μs pulse length. A range of* 30 *Km is achieved with a resolution of* 20 *m. The required precision of* 5 *milliradians is attained.*

The principles of the design of this radar are applicable to the design of radars for the detection of low-flying aircraft.

1. INTRODUCTION

La détection des objectifs mobiles au sol ou au voisinage immédiat du sol est rendue particulièrement difficile par la présence d'échos réflechis par les obstacles fixes environnants et que le pouvoir séparateur de l'antenne ne permet pas d'isoler.

En général, ces "échos fixes" ont des niveaux considérablement plus grands que les échos mobiles qu'on cherche à détecter.

Un rapport d'énergie réfléchie de 10,000 n'est pas rare, ce qui conduit à réaliser des radars ayant des facteurs de visibilité des échos mobiles dans les échos fixes de 40 db au moins.

Le problème qui a été posé par les Autorités Militaires au Laboratoire Central de Télécommunications était le suivant:

(a) Détection des objectifs mobiles au sol avec des portées de l'ordre de 30 km sur véhicule et 15 km sur homme isolé.

(b) Mesure des coordonnées polaires, gisement et distance de l'objectif détecté avec une précision respectivement de 5 millièmes et 40 mètres.
(c) Identification de la nature de ces objectifs (véhicules à roues, à chenilles, piéton isolé, groupe de piétons, etc. . . .).
(d) Identification des objectifs amis.

Le radar doit permettre la surveillance d'un secteur réglable avec une cadence d'information aussi élevée que possible, et de toute façon suffisante compte tenu des vitesses de déplacement des objectifs au sol. Il doit être facilement déplaçable et camouflable du fait qu'il opère près des lignes.

2. DETERMINATION DES CARACTERISTIQUES PRINCIPALES DU RADAR

2.1. *Spectre émis*

Les travaux sur la théorie de l'information appliquée au radar, développés si clairement dans les ouvrages de Woodward et de Carpentier, montrent

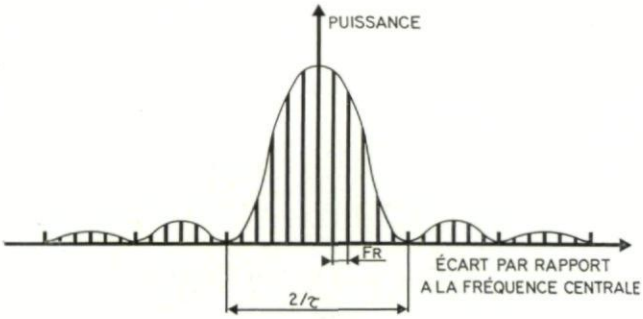


FIG. 1. *Spectre émis.*

que lorsque le spectre émis a une forme à peu pres rectangulaire, la discrimination en distance est donnée par la formule

$$d_{\min} = \frac{C}{2\Delta f}$$

C désignant la vitesse de la lumière et Δf la largeur du spectre émis. Par exemple, une discrimination en distance de 30 m conduit à un spectre de 8 MHz. Pour réaliser un tel spectre, il est apparu que l'un des moyens les mieux éprouvés était d'utiliser la modulation par impulsions qui convenait parfaitement aux magnétrons, tubes émetteurs sûrs, faciles à utiliser et de rendement élevé.

Le spectre de 8 MHz conduit à des impulsions de 0.25 μs tout-à-fait compatibles avec des magnétrons construits en grande série.

L'un des procédés de modulation les plus simples consiste donc à émettre des impulsions à cadence regulière, ce qui conduit à un spectre de raies du genre de celui de la figure 1. La largeur du spectre entre les deux premiers zéros est deux fois l'inverse de la durée d'impulsion: $2/\tau$; la distance entre raies consécutives est égale à la fréquence de répétition F_R.

2.2. *Spectre reçu*

2.2.1. *Spectre des échos fixes.* Le radar étant destiné à la surveillance, un temps limité θ est accordé à l'analyse d'un objectif particulier. Il en résulte que chacune des raies spectrales se trouve élargie comme l'indique la figure 2. La largeur des bandes élémentaires est $2/\theta$ c'est à-dire deux fois l'inverse du temps d'analyse.

En fait, si les échos fixes sont des arbres agités par le vent, un élargissement supplémentaire dû au mouvement des branches doit etre pris en considération.

2.2.2. *Spectre des échos mobiles.* Le spectre des échos mobiles, représenté en pointillé sur la figure 2, est indentique à celui des échos fixes, à la seule

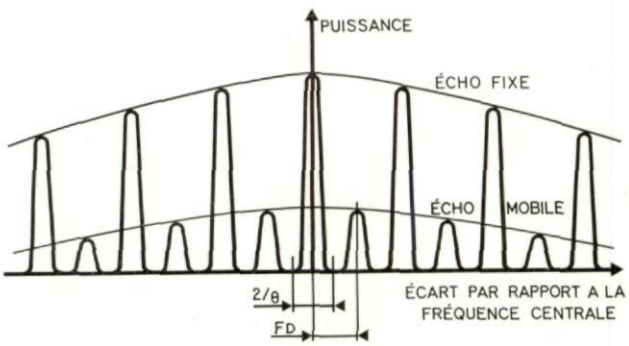


FIG. 2. *Spectre des échos fixes et mobiles.*

difference près qu'il a subi un décalage d'une quantité F_D par suite de l'effet Döppler.

2.2.3. *Separation des échos fixes et mobiles.* La condition nécessaire et suffisante pour effectuer par filtrage la séparation des échos fixes et mobiles est que leurs spectres soient nettement distincts.

Le "filtre adapté" correspondant à un écho mobile de vitesse radiale déterminée est un filtre dont la fonction de transfert en amplitude est identique au spectre de l'écho mobile considéré et dont la fonction de transfert en phase est l'inverse de celle du spectre. Si l'on désire "éliminer" les échos fixes, il faut donc que le spectre des échos fixes ne pénètre pas à l'intérieur du filtre adapté, et ceci avec un taux de réjection de 40 db. On est donc conduit à limiter la vitesse d'information à une valeur telle que les spectres d'échos fixes et mobiles n'empiètent pas les uns sur les autres.

2.3. *Emetteur-récepteur*

Les spectres qui ont été dessinés sont relatifs à un émetteur pour lequel les impulsions sont découpées dans une porteuse stable fournie par un pilote. Le magnétron étant malheureusement un auto-oscillateur, le spectre obtenu en pratique n'est pas un spectre de raies mais un spectre continu. Pour obtenir à la réception un spectre de raies, deux techniques peuvent être utilisées:

—celle du récepteur cohérent qui consiste à démoduler le signal reçu par une onde pilote mise en phase à chaque impulsion d'émission par le magnétron. Dans ce cas, il y a inversion entre le rôle du pilote et le rôle du tube émetteur par rapport à la solution de l'émetteur piloté. C'est ici l'émetteur qui détermine la phase du pilote et non pas le pilote qui détermine la phase de l'émetteur,

—celle du récepteur incohérent qui consiste à se servir des échos fixes comme onde de reférence pour démoduler les échos mobiles. Cette solution présente l'inconvénient de nécessiter la présence d'échos fixes pour pouvoir détecter un écho mobile. Par ailleurs, il est difficile par ce procédé de réaliser une sensibilité de détection constante des échos mobiles, lorsque des échos fixes sont superposés. Pour ces deux raisons, le récepteur cohérent a été préféré, malgré sa complexité plus grande.

2.4. *Longueur d'onde*

Le pouvoir discriminateur angulaire est donné par la formule:

$$\alpha_{\min} = \frac{70\lambda}{a}$$

$\alpha_{\min}$ représente l'angle minimum en degrés en dessous duquel deux objectifs sont confondus, λ la longueur d'onde et a la dimension de l'antenne.

Pour obtenir une discrimination en gisement de 1° tout en ayant une antenne de 2 mètres compatible avec un transport aisé, on est conduit à utiliser une longueur d'onde de l'ordre de 3 cm. Les longueurs d'onde plus courtes permettent des antennes moins encombrantes mais présentent les inconvénients suivants:

—absorption par l'atmosphère
—échos de pluie importants
—mauvais récepteurs
—technologie plus difficile
—perte de gain à la réception du fait de la surface réduite d'antenne pour un pouvoir discriminateur donné.

En fin de compte, la bande *X* a été choisie (8 500 à 9 600 MHz).

2.5. *Autres paramètres*

Les autres paramètres découlent de l'équation de portée du radar ou de considérations de bon sens.

Ils ont été groupés dans le tableau suivant:

—fréquence d'émission	bande X
—puissance de crête	40 kW
—fréquence de répétition	4 kHz
—durée d'impulsion	0.25 μs
—facteur de visibilité	40 db
—portée	30 km
—précision: gisement	$\pm$5 millièmes
distance	$\pm$20 millièmes
—vitesse radiale des objectifs	2 à 200 km/h

3. REALISATION PRATIQUE

3.1. *Realisation du filtre adapté*

La difficulté essentielle était celle de la réalisation du "filtre adapté," filtre donnant la sensibilité maximum pour une vitesse d'information considérée. Le récepteur cohérent dont on vient de parler brièvement présente déjà l'avantage de transformer le spectre hyperfréquence des échos fixes en spectre video fréquence centré autour de la fréquence zéro si l'on admet le concept des fréquences positives et négatives. L'information de signe de la vitesse radiale est d'ailleurs perdue du fait que l'on ne sait pas en pratique faire la différence entre les fréquences positives et négatives.

Malgré l'avantage du spectre video fréquence, la réalisation du filtre adapté n'en reste pas moins difficile car il est nécessaire en fait de construire un filtre présentant une suite de bandes passantes séparées par des bandes atténuées comme le montre la figure 2. Le nombre de bandes passantes est de l'ordre de 1000 et les coefficients de surtension pour chacune des bandes les plus éloignées sont tels que la construction du filtre par des méthodes conventionnelles a dû être écartée.

Dans le radar SDS, une approximation du filtre adapté a été réalisée comme l'indique la figure 3. On distingue sur cette figure:

—un "sélecteur," c'est-à-dire une porte à laquelle sont appliqués les signaux video fréquence à filtrer, et débloquée par une impulsion dite "créneau" décalée par rapport à l'impulsion de synchronisation d'une quantité connue et contrôlable,

—un "filtre Döppler," filtre passe-bande représentant la bande spectrale la plus proche de la fréquence zéro, du spectre d'écho mobile.

L'ensemble sélecteur et filtre Döppler constitue le filtre adapté, mais seulement pour une distance déterminée. Il est en effet possible de voir, en plaçant un générateur à l'entrée A du système et un voltmètre à la sortie B, que la courbe de réponse est constituée par une suite de bandes passantes et atténuées comme l'indique la figure 2. Mais il est aussi évident que si un écho mobile n'est pas en coïncidence avec le créneau, aucun signal ne sort du dispositif.

Le filtre adapté valable pour toutes les distances serait donc constitué par une batterie de dispositifs identiques à celui de la figure 3 et alimentés par des créneaux de distance croissante. La vitesse d'information étant dans ce cas très surabondante, on s'est contenté, dans le radar SDS, de construire 5 chaînes de ce type. Pour simplifier l'équipement, le filtre Döppler couvre en outre la totalité de la bande des objectifs mobiles possibles, ce qui conduit à une perte de sensibilité qui n'est pas en fait gênante compte tenu ici encore des marges de sécurité importantes qui ont été prises. Deux types de

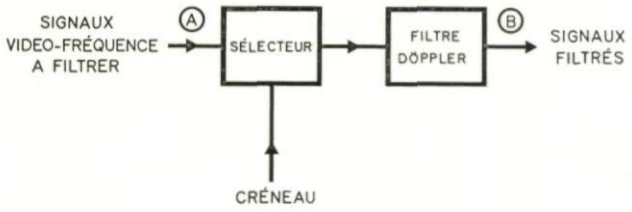


FIG. 3. *Système de filtrage*

filtres Döppler existent suivant qu'il s'agit de détecter des échos rapides ou des échos lents (piétons):

—pour les échos rapides, le filtre couvre une bande de fréquence comprise entre 120 Hz et $FR/2$ soit 2000 Hz. Cela permet de détecter avec une sensibilité constante les objectifs dont les vitesses sont comprises entre 6 km/h et 200 km/h environ. Après une zone de vitesses aveugles de 12 km/h d'amplitude, une nouvelle zone de détection apparaît, et ainsi de suite,

—pour les échos lents, le filtre est à bande plus étroite et permet de couvrir de façon continue la gamme de vitesses comprise entre 2 et 9 km/h.

3.2. *Schéma général*

La figure 4 représente le schéma général qui comporte deux parties:

—un émetteur-récepteur cohérent,

—un ensemble d'élimination d'échos fixes.

L'émetteur-récepteur cohérent a déjà été décrit dans la littérature.[1]

A la différence des émetteurs-récepteurs pilotés, il comporte un auto-oscillateur à l'émission. Afin d'avoir une référence de phase, un "oscillateur cohérent" est mis en route en même temps que le magnétron qui constitue l'auto-oscillateur. Une faible partie de la puissance du magnétron est injectée après changement de fréquence dans le circuit oscillant de l'oscillateur cohérent qui démarre ainsi sur la phase du magnétron et conserve la mémoire de cette phase pendant le retour des échos.

La sortie de l'oscillateur cohérent est utilisée comme onde de référence pour démoduler les signaux moyenne fréquence en provenance des échos.

Si la relation de phase entre l'onde de référence et le signal est constante, il s'agit d'un écho fixe.

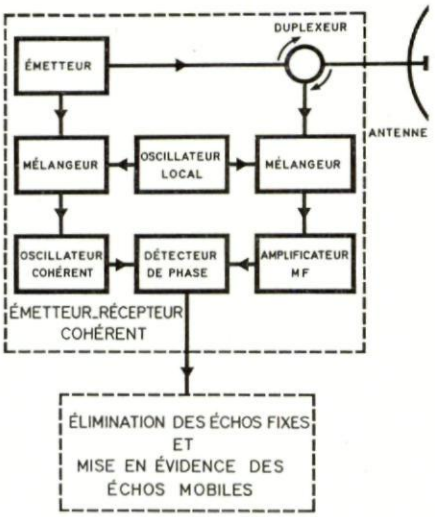


FIG. 4. *Diagramme général.*

Si la phase glisse linéairement en fonction du temps d'un balayage au suivant, il s'agit d'un écho mobile.

La figure 5 représente la succession des échos fixes et mobiles à la sortie du récepteur.

La figure 6 représente l'aspect des signaux sortant du récepteur, l'impulsion d'émission servant à synchroniser l'oscilloscope.

3.3. *Localisation de l'objectif, Identification*

La figure 7 montre comment s'effectue la localisation d'un objectif. Le créneau qui sélectionne un objectif sur l'échelle des distances est contrôlé avec précision à l'aide d'un système électro-mécanique commandé par une manivelle. Une autre manivelle commande le mouvement de l'antenne. Un

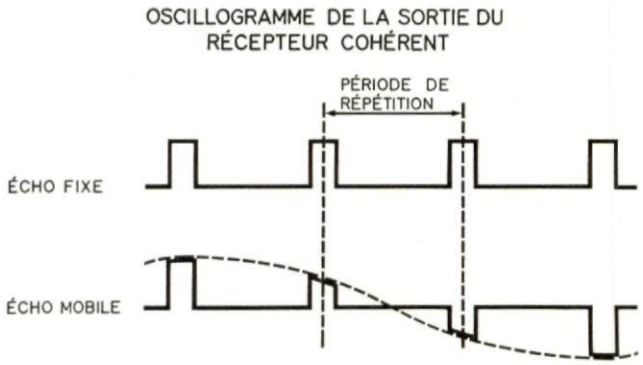


FIG. 5. *Oscillogramme de la sortie du récepteur cohérent.*

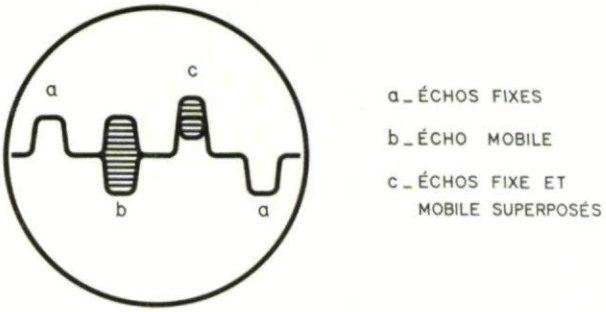


FIG. 6. *Oscillogramme de la sortie du récepteur cohérent.*

haut-parleur est placé à sortie du filtre Döppler. En réglant les deux manivelles pour obtenir le niveau sonore maximum, on lit directement sur des cadrans le gisement et la distance.

Le son Döppler est en général complexe car le radar reçoit simultanément plusieurs vitesses radiales différentes: mouvement des bras, des jambes et du corps d'un homme marchant, mouvement des éléments de chenilles d'un char, etc. L'oreille constitue un organe d'analyse extrêmement efficace pour effectuer la discrimination entre les différents types d'objectifs.

3.4. *Fonctionnement en surveillance*

Pour assurer une surveillance continue d'un secteur, on pourrait imaginer que l'opérateur tourne lentement sa manivelle de distance pour passer en revue les échos successifs à un même gisement. Puis, après avoir exploré la zone de distance qui l'intéresse, il décalerait son antenne d'une quantité égale à la largeur du faisceau et recommencerait une nouvelle exploration, et ainsi de suite. En fait, ces mouvements sont effectués automatiquement en distance et en gisement. L'écoute sonore est remplacée par une observation oscilloscopique en type B, c'est-à-dire:

—horizontalement, le gisement,
—verticalement, le distance.

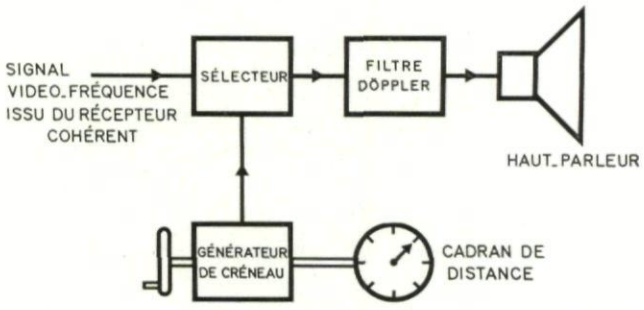


FIG. 7. *Diagramme de la mesure de distance.*

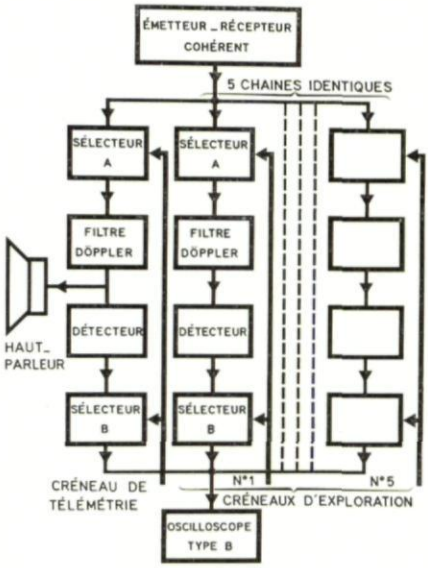


FIG. 8. *Diagramme d'ensemble.*

A des fins de simplification, il a été admis que la zone de distance surveillée était de 5 km, l'origine de cette zone étant réglable de façon continue par l'opérateur.

Pour augmenter la vitesse d'information, cinq chaînes identiques sont montées en parallèle comme le montre la figure 8. Une sixième chaîne permet d'effectuer la localisation.

Les cinq créneaux d'exploration automatique sont décalés d'un kilomètre l'un par rapport à l'autre et défilent sur l'oscilloscope type B comme le montre la figure 9. Il y a lieu de noter que pour restituer l'information de distance sur l'oscilloscope, les signaux de sortie des filtres Döppler sont détectés puis remodulés dans les sélecteurs B par les créneaux.

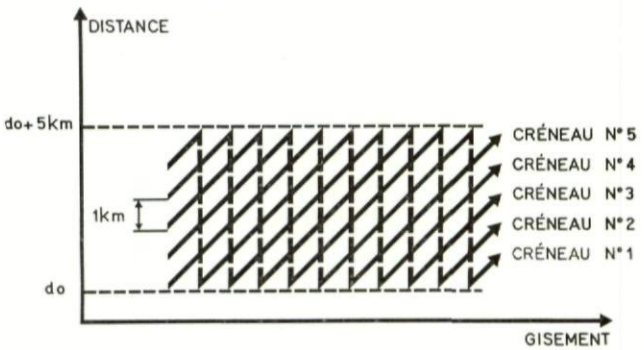


FIG. 9. *Filtrage des échos fixes dans un radar à ligne a retard.*

3.5. *Mode d'emploi du radar*

L'opérateur place d'abord son radar en surveillance, choisit les limites du balayage secteur et la zone de distance surveillée. L'antenne explore le terrain en gisement tandis que les créneaux défilent sur la zone de distance choisie. Lorsqu'un point lumineux apparaît, l'opérateur arrête le mouvement de l'antenne qu'il contrôle alors à la main au moyen d'une télécommande; il cherche à amener un marquer sur le point lumineux que la rémanence de l'écran maintient visible pendant plusieurs secondes. Il entend alors le son Döppler qu'il cherche à rendre maximum. Les cadrans de gisement et de distance lui donnent les coordonnées du but et le son entendu lui permet de connaître la nature de l'objectif.

3.6. *Identification des objectifs amis*

La firme Electronique Marcel Dassault a créé pour les radars de surveillance du sol, une balise répondeuse portative recevant les signaux et les retransmettant dans la bande VHF. Un adaptateur permet alors de visualiser les réponses sur l'indicateur du radar tandis qu'un code Morse remplace le son Döppler pour l'identification sonore.

4. COMPARAISON AVEC LES RADARS A ELIMINATION D'ECHOS FIXES CLASSIQUES

Ces radars sont constitués en général par un émetteur-récepteur cohérent suivi d'un dispositif d'élimination d'échos fixes composé d'une ligne à retard

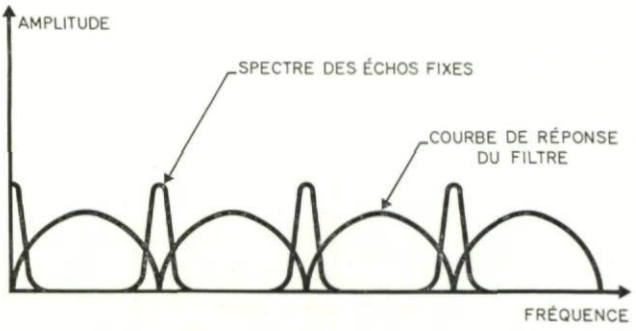


FIG. 6.

et d'un circuit de soustraction dans lequel est élaborée la différence entre deux sondages successifs. Ce dispositif constitue un filtre à bandes éliminées multiples dont la fonction de transfert est représentée Fig. 10. Il est facile d'observer sur la figure que le spectre des échos fixes pénètre de façon important dans la bande passante du filtre; il en résulte que le facteur de visibilité est limité à une valeur assez basse, en général incompatible avec la détection des objectifs terrestres.

5. DETECTION DES AVIONS VOLANT BAS

Le facteur de visibilité élevé du radar SDS permet d'envisager favorablement l'utilisation des principes mis en jeu pour la réalisation de radars de détection d'avions volant bas.

Les appareils tels qu'ils sont réalisés actuellement n'ont pas une vitesse d'information suffisante, mais il a été vu qu'en fait il était possible de l'augmenter considérablement en accroissant le nombre de créneaux d'exploration. La sensibilité peut également être accrue dans un rapport important en multipliant les filtres Döppler.

De nouveaux procédés, faisant appel à des techniques digitales, sont actuellement à l'étude au L.C.T. en liaison avec la Direction des Recherches et Moyens d'Essais. Ces procédés permettront dans un avenir très prochain de tirer la quintessence des radars SDS et en particulier de résoudre de façon satisfaisante le problème de detection des avions volant bas.

6. RESULTATS OBTENUS

Les appareils SDS ont été expérimentés à l'état de prototypes en France, en Algérie, aux Etats-Unis. Ils ont été mis en service dans la troupe et ont été utilisés avec succès sur les barrages tunisiens et marocains.

Une expérience intéressante a pu être ainsi acquise par:

—la Section Technique de l'Armée,
—les Unités combattantes,
—les Services d'Evaluation du matériel électronique de l'Armée américaine.

Les portées calculées ont été confirmeés:
—15 km sur piéton isolé et
—35 km sur véhicule.

Dans certains cas, des échos de piéton ont même été recueillis au delà de 20 km.

L'identification sonore s'est révélée très efficace et les précisions contractuelles de 5 millièmes en gisement et 40 mètres en distance ont été largement tenues et meme souvent dépassées.

Le matériel s'est révélé résistant et s'est bien comporté en Algérie ou les transports quotidiens et le climat rendaient le service particulièrement dur.

7. REMERCIEMENTS

L'auteur tient à remercier tout particulièrement:
—les Ingénieurs Militaires de la SEFT qui ont su, grâce à la continuité et la compétence de leurs directives, permettre au L.C.T. de mener à bien une étude s'avérant difficile au départ,
—les Officiers de la STA qui, grâce à leur expérience, ont permis de rendre le matériel opérationnel,

—l'Armée française qui a su s'adapter aux radars et les exploiter avec efficacité,
—les Autorités américaines, qui ont fourni une contribution financière au moment de l'étude et ont conduit avec efficacité des campagnes d'essais,
—le Service Cinématographique de l'Armée, qui a realisé un film remarquable sur le Radar de Surveillance SDS,
—enfin les Ingénieurs, Techniciens, Dessinateurs et Ouvriers du Laboratoire Central de Télécommunications qui, grâce à leur esprit d'équipe, ont contribué au succès final.

REFERENCE

1. RIDENOUR, L. N. *Radar System Engineering* M.I.T. Series Vol. 1, McGraw-Hill 1947.

CHAPTER 11

SIDEWAYS LOOKING AIRBORNE RADAR

J. HIGHCOCK
Royal Radar Establishment, Malvern, England

1. INTRODUCTION

The development of Sideways Looking Radar started at R.R.E. around 1950 and since then much has been done to realise the inherent capabilities of the non-coherent technique. Basically the concept is very simple; a long aerial is mounted along the side of the aircraft so that it looks at right angles to the flight path (see Fig. 1).

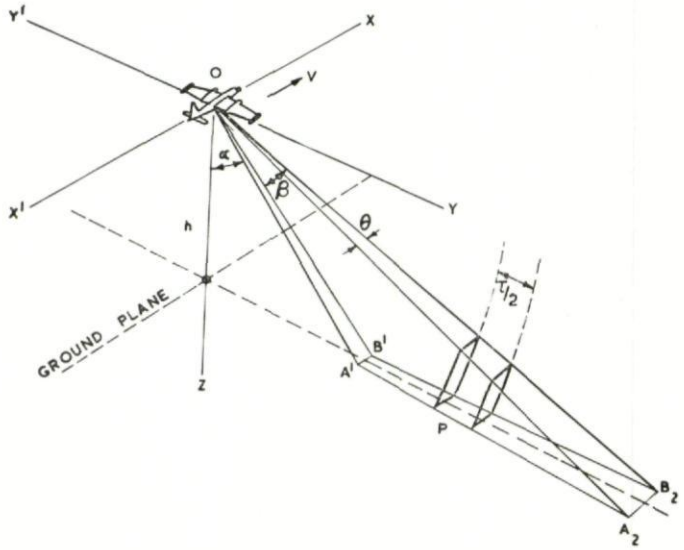


FIG. 1. *Sideways looking radar principle.*

The azimuth or along track beamwidth of the aerial is inversely proportional to the length of the aerial, which is also arranged to have a vertical beamwidth of anything up to 45°. The signal returns from each scan are displayed as intensity modulation (Z scan) on a cathode ray tube (C.R.T.) (see Fig. 2), the resulting line being projected on to a photographic film, which moves continuously at a rate proportional to the speed of the aircraft. The picture produced is of a strip of ground to one side of the aircraft and parallel to the flight path, the resolution of the map being a function of the beamwidth along track and the pulse length across track.

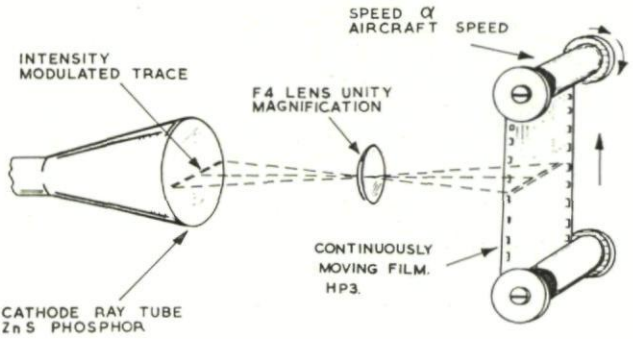


FIG. 2. *Elements of photographic recorder.*

2. AERIALS

2.1. *General*

The aerial which has been mostly used in the development of S.L.R. is the slotted waveguide linear array (see Fig. 3) which lends itself very readily to the generation of horizontally polarised radiation. It has a certain basic simplicity in that the power is fed into one end and leaks out of the waveguide from slots cut into the narrow face of the guide. Any power remaining beyond the last slot is merely absorbed into a load at the far end of the waveguide. The radiation from the slots is usually arranged to be uniform in phase so that a plane wavefront is created, effectively focussed at infinity. The amplitude distribution across the horizontal aperture is designed to follow one of the accepted forms either to minimise the side lobes or to maximise the gain, by varying the coupling of the slots into the guide as a function of distance along the guide or of the field strength within the guide. The vertical aperture is generated by a horn which can be asymmetrical when used alone or symmetrical with a lens, and in the throat of which the guide is located. Shaping of the vertical pattern can be achieved in a variety of ways to produce a cosecant^2 law, which can be optimised for a particular altitude of operation.

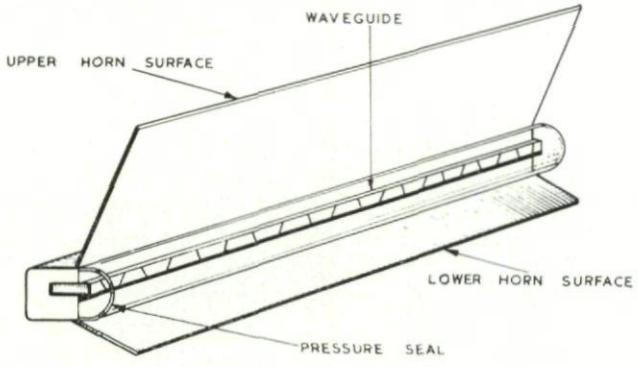


FIG. 3. *Linear array.*

This type of aerial is dispersive to a mild degree and when used with very short pulses certain precautions may have to be taken to prevent broadening of the beam. The frequency sensitive properties can be exploited in designing a yaw corrected aerial which can be quite a useful feature in an unstable aircraft. Since mechanical tolerances are quite severe it is necessary to make the aerial as rigid as is practicable by mounting it on a substantial metal beam; the resulting arrangement can be quite heavy and if mounted on the aircraft at one point only, it has the desirable property of being somewhat insensitive to minor aircraft motions.

It is normal to transmit and receive on the same aerial, so that the two way pattern is merely the (one way pattern)^2.

2.2. *Shape of Vertical Pattern*

The vertical pattern is shaped in an attempt to equalise the signal returns from all ranges (see Fig. 4). The angle β over which the aerial looks is arranged to cover QT, a strip of ground maybe 10 miles wide, where QOP

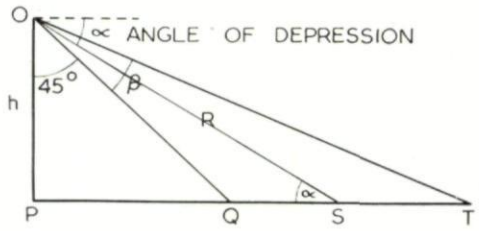


FIG. 4. *Geometry determining vertical pattern.*

the minimum angle of incidence is not less than 45°. The reason for this is simply that at angles less than 45° the difference between successive slant ranges is very small so that the radar has little or no effective range discrimination. It should also be noted that the radar is conscious of slant range R and not ground range PS, so that with a linear time base on the recording C.R.T. some distortion of a plan range picture occurs.

The radiated power reaching a point reflector or scatterer at S is inversely proportional to R^2, as is the power reaching the aerial O from S on the return journey. For equal reflectors which are not aspect conscious the one way pattern required is proportional to cosecant^2 α, where α is the angle of depression or complement of the angle of incidence. The beam shaping required for equal ground returns also depends on the reflection coefficient which decreases with angle of depression and on the area of ground filling the beam. The instantaneous area of ground which reflects energy back to the radar is proportional to the pulse length and the beamwidth so that the effective reflecting area is equal to

$$R \cdot \frac{\tau}{2} \cdot \theta \cdot c \cdot \sigma$$

where θ is the azimuth beamwidth, τ is the pulse length, c is the velocity of EM waves, and σ is the reflection coefficient.

It should be noted that σ is defined as the ratio of the energy scattered isotropically by the surface to that incident on the surface. If this expression is substituted in the radar equation, the usual R^4 term in the denominator becomes R^3, but is balanced by σ which appears in the numerator. The law of variation of σ with the angle of incidence is dependent on the precise value of the angle of incidence and on the type of terrain which is being observed.

The waveguide is situated in the throat of the horn, so that the effective vertical aperture is defined by the flared ends of the horn. An asymmetric

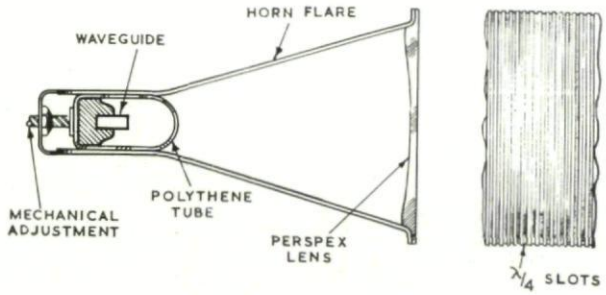


FIG. 5. *Typical aerial.*

horn fed off axis can be used on its own to shape the beam, either for low altitude or high altitude operation. Alternatively a dielectric (often polystyrene) lens can be introduced across the exit of a symmetrical horn the particular virtues of a lens as compared with a mirror being its ease of manufacture and its more reasonable mechanical tolerances, although it can have the disadvantage of being relatively heavy when the aerial is long. The deviation through the lens as a function of the primary feed angle can be estimated by equating the angular energy distribution of the primary feed and the angular energy distribution required of the radiated pattern. If the front surface of the lens is plane the rear surface will have a profile as shown in Fig. 5. The front surface of the lens is usually matched to free space by introducing quarter wave slots cut in the dielectric.

2.3. *Horizontal or Azimuth Pattern*

The radiating slots are cut in the vertical narrow face of the waveguide and to minimise the frequency sensitivity the distance between slots is made slightly greater than $\lambda g/2$, which makes the aerial non-resonant and introduces a small angular squint of about 2° or 3°. The array will still be frequency sensitive to some degree (see Fig. 6), so that changes of magnetron frequency during life will cause the squint angle to change slightly. Edge slots are used since they can be located and cut quite accurately and the

resulting radiation is horizontally polarised. Since the slots are inclined to the vertical the E vector is not quite horizontal, but because the vertical components are cancelled out by being successively out of phase, the radiation is horizontally polarised. A certain amount of power is lost in this way, which increases with the inclination of the slot. The effect can be allowed for by increasing still further the inclination and conductance of the slots towards the end of the array. Horizontal polarisation is preferable to vertical by virtue of the improved contrast obtainable over the wide range of natural and artificial features normally surveyed.

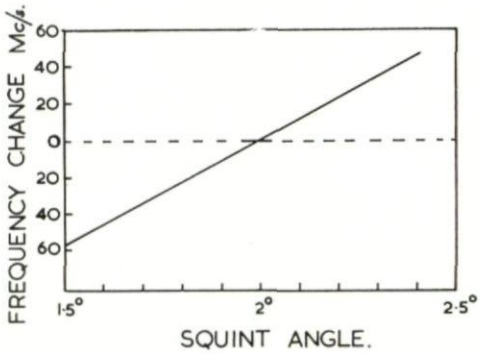


FIG. 6. *Aerial frequency sensitivity.*

In an X-band array the tolerance on the spacing of the slots is about 0.0005 ins., but a cumulative error is avoided by tolerancing the two ends of the aerial to about 0.005 ins.

2.4. *Other Properties*

It takes a finite time for the pulse of energy to fill the guide and in this time the power radiated is low and the effective beamwidth is large, becoming rapidly narrower until the energy fills the guide. It is necessary for the time to fill and empty the guide to be a small proportion of the pulse length. The problem can be looked at in another way by considering the dispersive properties of the guide. It is necessary for the bandwidth of the pulse to be small compared with the frequency sensitivity of the guide so that frequency components of any magnitude within the pulse should not cause widening of the beam. This problem does not become serious until pulse lengths of less than about 50 nanoseconds are used, unless extremely long aerials are contemplated.

3. RECORDING AND VIDEO INTEGRATION

3.1. *General*

As indicated in Fig. 2, after demodulation the signals are applied to a C.R.T. as intensity modulation of a linear scan (Z scan) and are then

projected onto a continuously moving film to produce a plan range map. A good deal of the effectiveness of the S.L.R. technique is contained in this relatively simple process, one of the most important elements of which is the integration due to the super-position of successive scans on to the photographic film. Also of great importance are the performances of the C.R.T., lens and film which must have resolutions capable of handling the inherent resolution of the radar signals.

3.2. *Video Integration*

The quality of the map produced by a S.L.R. is affected very appreciably by the signal to noise of the recorded signals and by their relative contrast. It is possible to improve the signal to noise ratio by post detector or video integration, the process being the natural result of superimposing on a photographic film many scans or returns from the same effective strip of ground. The improvement can be shown to be proportional to $\sqrt{n}$ where n is the number of scans involved in integration, a result which is typical of post detector integrating circuits. The contrast of the picture is independent of the integration process, but depends upon the effective gamma of the C.R.T. and film.

In a typical X-band S.L.R. with a beamwidth of 0.4°, several hundred signals will be received from a point reflector as it passes through the beam, even at short ranges. At long ranges the figure can rise to thousands of pulses. A possible improvement of $\sqrt{1000}$ in the signal to noise ratio is obviously of great importance, and it is important to have a clear understanding of how the process occurs in order to design the radar to take full advantage of it.

The effect of integration is obviously of most interest where small signals are concerned and for which it may be assumed that the receiver is linear up to the second detector. A square law second detector is a reasonable assumption for small signals i.e. $y \propto v^2$ where y is the output voltage and v the input voltage.

The video signals from the second detector are amplified in a linear amplifier and are applied to the grid of a cathode ray tube having a luminous intensity versus grid drive law, which can vary with the type of tube used and the operating point chosen, thus $I \propto (y - y_0)^P$ where I is the intensity of the C.R.T. spot, y the drive signal and y_0 the cut off voltage.

The light signals registered on the screen of the C.R.T. are collected by a lens and projected as a real image on to the photographic emulsion. The law of density D versus exposure E can be expressed in the form:

$$D = \gamma \log_{10} \frac{E}{E_c} \text{ for } E > E_c$$

where E is the total exposure of the film and E_c is a constant. The gamma (γ) of the film refers to the slope of the linear part of the curve, and this can be controlled to some extent in the development process. It is also dependent on the type of film selected.

The output signal of importance is the resultant density of the photographic film, and it is convenient to measure the effectiveness of the recorder at this point.

In Fig. 7 it can be seen that a mean density D_0 occurs when noise of variance σ^2 is present, and when a signal is applied and noise of about the same variance is present at the same time, the mean density level rises to D. The detectability of the impressed signal will depend on the magnitude of the mean level D when the signal is present, to the mean level D_0 when only noise is present. It is possible to define the output signal to noise ratio of the film as $S = \dfrac{D - D_0}{\sigma}$. This is not the sole criterion of performance however since the signal will only be detectable if the contrast between D and the

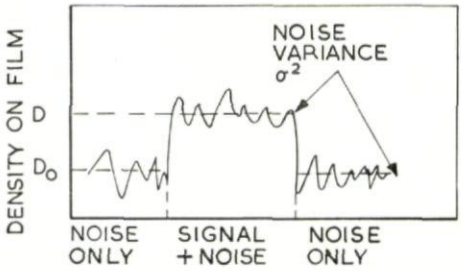


FIG. 7. *Contrast and signal to noise ratio.*

background is adequate. Although the contrast can be expressed as $C = D - D_0$, its useful limit is determined subjectively and cannot be stated with precision. It is necessary for both S and C to be sufficiently large for good quality recording and for the ultimate realisation of the high resolution properties of the radar.

It may be seen that

$$C \simeq \gamma \log_e \frac{I - I_0}{I_0}$$

and

$$S \simeq \sqrt{n} \cdot \frac{I - I_0}{\sigma_g}$$

where I and I_0 are the luminous intensities of the C.R.T. with signal and noise and with the noise alone respectively, n is the number of scans integrated and σ_g is the variance of the noise luminous intensity. It should be noted that S and C are independent of the lens stop and optical efficiency, although it is necessary for sufficient light to reach the film to prevent reciprocity failure.

3.3. *Effect of Spot Diameter*

It cannot necessarily be assumed that the integration of signal and noise always occur for the same time, as both depend on the C.R.T. spot size and

shape, the light output being dependent on the convolution of the grid waveform and the spot shape and intensity distribution. In addition it is necessary to take into account the time during which the film traverses a spot width and the time available for integration. If these are unequal a condition arises where the signal is integrated for n_1 scans and the noise for n_2 scans, so that

$$S = A \cdot \frac{n_1}{\sqrt{n_2}} \theta$$

and

$$C = \gamma \log_e \cdot A \cdot \frac{n_1}{n_2} \cdot \phi$$

where A is the signal to noise ratio from the second detector, and θ and ϕ are constants of proportionality associated with the signal to noise and contrast respectively.

The important factors to note here are that for a fixed value of n_2 both S and C increase linearly with n_1, until $n_1 = n_2$. The spot size should be not larger than the equivalent pulse length, and to prevent any increase of effective pulse length the optimum receiver bandwidth should be used. It is obviously important to make sure that the time to traverse a spot diameter is not longer than the time for which a signal is present. On the other hand if it is shorter, only a portion of the available pulses will be integrated.

The constant of proportionality θ is dependent on the power law P connecting the C.R.T. light output with the grid drive and it is also a function of the grid bias voltage. The value of θ is always less than unity, and it decreases with increasing values of p and bias. For a square law tube, which is fairly normal, and with a normalised negative bias of unity (i.e. normalised with respect to the noise power), θ is approximately 0.95, so that

$$S \simeq A \cdot \frac{n_1}{\sqrt{n_2}}$$

If $n_1 = n_2$, $S \simeq A \cdot \sqrt{n}$. The contrast increases with bias and p, since ϕ is proportional to the input signal to noise ratio. For the same negative bias setting of unity, a square law tube gives a value of 1.2 for ϕ. Adjustment of the bias permits an effective means of obtaining a compromise between signal to noise and contrast, so that p will vary between 0.6 and 2 for values of θ around 0.95.

If a minimum detectable density difference of 0.06 is assumed and $\gamma = 1$, for 1000 integrated scans and a square law tube, $\theta = 0.95$ and $\phi = 1.35$. With parameters such as these it is feasible to detect signals in the range from 5 to 10 db below noise if an output signal to noise ratio of 10 : 1 is assumed to be necessary for detection.

3.4. *Resolution of C.R.T.*

The inherent resolution and the mapping range of the radar are such that great demands are usually made of the recording system. For instance

a radar with a pulse of 0.2 μs mapping over a range of ten miles requires 600 range increments or elements to be recorded. If both sides of the aircraft are being mapped simultaneously, which is often the case, twenty miles has to be recorded with a resolution of 100 ft., using perhaps a 70 mm C.R.T. and film.

It has been shown that it is advantageous for the C.R.T. spot size to be approximately equal to the equivalent pulse length of the radar. In general the C.R.T. spot size and pulse length will be less than the equivalent linear beamwidth of the aerial except at short ranges, but this does not affect the criterion of equating pulse length and spot size. The spatial resolution of which a C.R.T. spot is capable is a function of its size and shape and of the light intensity distribution within it.

The spatial frequency spectrum of the C.R.T. spot distribution is the spatial frequency response of the C.R.T. and is analogous to the frequency response of an electrical network. The spatial frequency response is a maximum at zero frequency and falls off with increasing frequency, becoming effectively zero when $2/(2\pi f)$ is less than the spot size. It is usually normalised in terms of the zero frequency response (see Fig. 8).

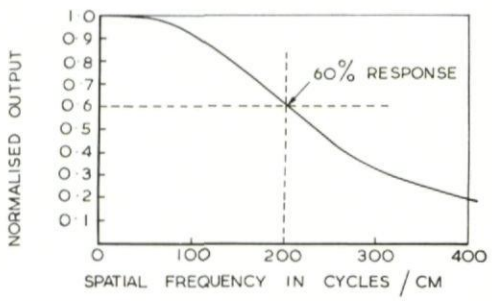


FIG. 8. *Typical spatial frequency response of C.R.T.*

The spatial frequency response of a C.R.T. has a typical low pass characteristic, which is approximately half-Gaussian in shape. Since the spatial frequency responses of the other elements of the display, the spatial frequency responses of the other elements of the display, i.e. lens and film are also half-Gaussian in shape, albeit with different effective "pass-bands," the frequency response of the whole display is half-Gaussian and the effective pass band is obtained by taking the R.M.S. value of the individual pass bands.

Since measurements of the performance of C.R.T.'s were started a few years ago, significant improvements in resolution have occurred, until at the present time it is possible to produce a C.R.T. which is as good as a lens and film.

3.5. *C.R.T. Characteristics*

The C.R.T. has to produce an adequately bright and small spot under all operating conditions. Much work has been done on electromagnetic tubes to

improve their performance to that available today. Electron guns now have a small crossover where the beam passes through the modulating region. If the crossover is regarded as the object of the electron lens, the latter "magnifies" it without introducing significant aberration by using a small aperture in the anode to limit the cross sectional area of the beam. External manipulation of the beam along the axis of the lens by field correcting solenoids results in a minimum of astigmatism and coma distortion.

Fine grain settled phosphors carefully deposited to reduce the occurrence of pin holes and agglomerates can be made to produce adequate light output with spot sizes of about 10 microns. The noisiness of such screens, due to the uniformity of deposition, is very low, a factor of some importance where picture generation is concerned. A figure of 5% in the variation of intensity across the screen is now fairly typical.

An electromagnetic tube requires a very stable E.H.T. voltage, a change of 1% causing deformation of the spot and a reduction in resolution of about 50%. The acceptable tolerance on the E.H.T., which may be 10 to 15 kV is 0.25%, and 0.5% on any superimposed ripple. Quite satisfactory results can be obtained by using corona tube stabilisers operated from a smoothed d.c. supply. The focus current also has to be kept constant to about 0.1%, a change of 0.5% causing a reduction of 50% in the resolution. The condition can be met by using low impedance focus coils supplied from a transistor current stabiliser.

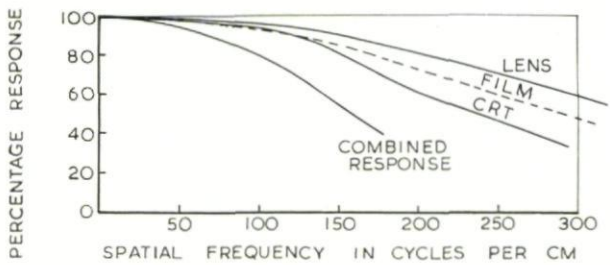


FIG. 9. *Display system frequency response.*

As a standard of comparison a fine grain zinc sulphide screen with adequate light output will have a spatial resolution of 200 cycles per cm measured at the point at which the response has fallen to 60% of its d.c. value. Used with such a tube would be a F/4 lens of 300 cycles per cm and HP3 film of 250 cycles per cm, the three components giving a combined 60% response of about 140 cycles per cm (see Fig. 9).

3.6. *Overall Resolution of Radar*

The effective resolution of a S.L.R., modified by the characteristics of the receiver and the display is determined by the pulse length and the beamwidth. If a perfect recorder is assumed it can be shown that the inherent across track and along track resolution as they would appear on the display are

Gaussian in form, so that a figure defining overall resolution can be obtained by taking the R.M.S. value of those components combined with those of the display.

The across track resolution of the radar is determined by the spectral width of the transmitted pulse modified by the bandwidth of the receiver. A transmitted pulse may have a shape not far off rectangular, but more probably nearer to a sine squared function. The pulse emerging from the receiver will approximate to a Gaussian shape, the half spectral width or $(\text{variance})^{\frac{1}{2}}$ corresponding to the 60% power points, defining the effective resolution of the transmitter-receiver combination. If the temporal spectral width is F_B, it can be converted into a spatial frequency $f_s = \dfrac{2F_B}{c}$ cycles per unit length across track where c is the velocity of E.M. waves. The factor 2 arises since the effective pulse length is half the actual pulse length. If the width of the track is G, then $f_s = \dfrac{2F_BG}{c}$ cycles per track width or scan length.

The inherent along track resolution of the radar is determined by the shape of the beam and the range. As the aerial scans the ground the signal returns are modulated by the square of the polar diagram and this function is converted by the display, in terms of the display coordinates, into a similar function $G(s)$ in a direction perpendicular to that representing the beam. If the display is assumed perfect the function $G(s)$ will define the frequency response of the aerial polar diagram in the relevant coordinates.

It can be shown that the half spectral width of this aperture distribution at a slant range R is

$$\frac{a}{R\lambda}\sqrt{\frac{2}{3}} \text{ per unit length across track,}$$

where $2a$ is the aerial aperture.

The along track resolution deteriorates with range so that if a radar of pulse length 0.1 μsec and beamwidth 0.006 radians is contemplated, the along track resolution will vary between 750 cycles per 70 mm trace length at 5000 ft to 125 cycles per trace at 60,000 ft, the trace representing a 5 mile range interval. The corresponding figure for the across track resolution is constant at 400 cycles per 70 mm trace. The recorder can have a resolution of about 1000 cycles per 70 mm trace. The overall resolutions along and across track can be calculated, and it should be noted that the display causes very little degradation of the inherent radar resolution.

3.7. *Camera and Film*

When all has been done to equalise the signals from effective reflecting areas of the same magnitude at all ranges, the dynamic range of the input signals can still be very large, in the range of 40 to 50 db. The dynamic range of photographic film is limited, so that the range of signals capable of being handled by a square law C.R.T. and film combination will be between 15 and 20 db, assuming first a low chemical fog density level of 0.2 or 0.3,

and secondly that the film is developed to a gamma of unity. It is standard practice to employ a logarithmic I.F. amplifier which is linear for small signals, but which compresses larger signal to noise ratios in the range 10 to 50 db to the range 10 to 20 db.

Variations of aircraft speed will cause more or fewer scans per unit length travelled and since the film is pulled through the camera at a rate proportional to the ground speed of the aircraft, the brightness of the picture will fluctuate unless some form of compensation is applied such as variation of the lens stop or brightness of the C.R.T. trace. The former is preferable since varying the C.R.T. bias voltage will change the law of the tube and hence the picture density (or overall gamma) scale. Steps can be taken to ensure constant "base" brilliance (for zero signal input) of the C.R.T. by monitoring the light output with a photocell in the dead time immediately following the scan whilst the spot is stationary.

The camera mechanism has to be smooth and reliable in operation and much work has been done over the years to improve designs to their present satisfactory state. The use of sprocket driven film is to be avoided since striations are very readily produced on a film moving as slowly as an inch or so per minute. One of the best methods is to introduce loops of film between the feed and take up sprockets and the capstan, so that the latter tends to be isolated from jerks caused by the sprocket. If the capstan is driven it can be made to move the film by friction and it is only necessary to ensure that sprockets and capstan move effectively at the same rate in order to maintain the loops in the film.

3.8 *Radar Mapping*

The coordinates of the map produced by a S.L.R. are slant range along the beam and plan range along track, and distortions will be produced by variations of pitch, yaw, drift and aerial squint. In order to produce an accurate plan range picture the C.R.T. recorder would have to use a hyperbolic time base which is rather difficult to provide, but it is generally considered preferable to use a linear time base and to record along the side of the film the attitude and position of the aircraft, so that accurate positional information can be obtained by subsequent computation. The geometry is indicated in Fig. 10.

In the horizontal plane the aerial is displaced by an angle ψ which is made up of yaw and drift. In the vertical plane the aerial is displaced by an angle θ which is due to the pitch of the aircraft. In addition the aerial beam squints from the normal to the aerial by an angle α. The effect of squint is to cause the aerial beam to look along the surface of a cone of half angle $(90 - \alpha)$, the axis of the cone being the aerial itself (see Fig. 11). The intersection of this conical beam with the horizontal plane defined by the ground is a section of a hyperbola, instead of the straight line produced by an orthogonal beam. It can be shown that negligible error results, except at very long ranges, if the beam is assumed to be flat and the intersection QP is therefore assumed to be a straight line.

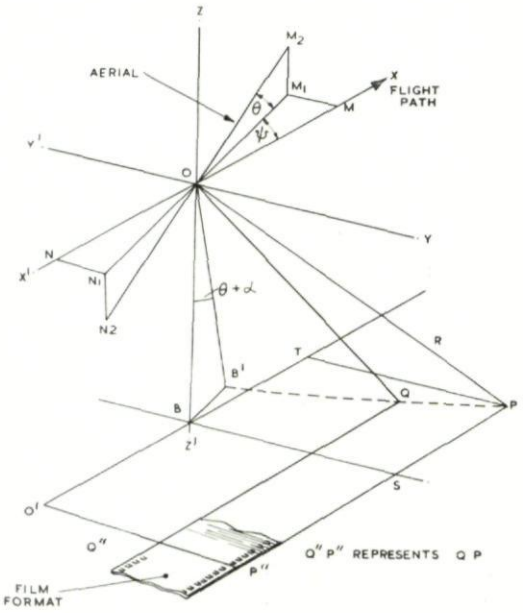


FIG. 10. *Sideways looking radar map geometry.*

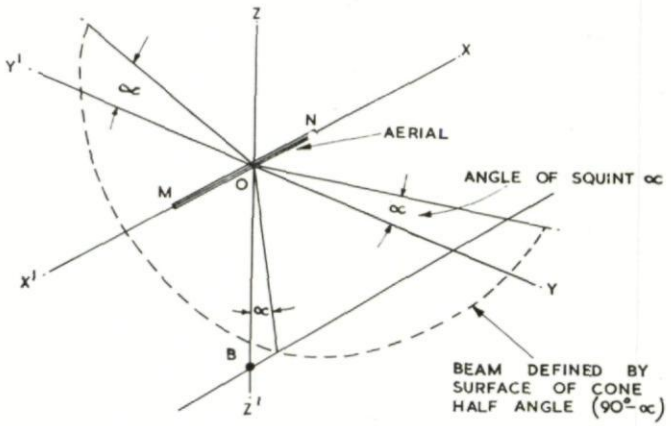


FIG. 11. *Effects of aerial squint.*

Figure 12 refers the point P to axes defined by the heading of the aircraft, so that only the pitch angle θ and the squint angle α need be considered. The squinted beam $ODGP$ intercepts the ground plane through P in the line DP, which although part of a hyperbola, can be assumed to be a straight line.

The required coordinates are $x_1 = PS_1, y_1 = PT_1$ and $z_1 = OB = H$

$$\therefore\ x_1 = R \sin \alpha \sec \theta + H \tan \theta \tag{1}$$

$$y_1 = PT_1 = (R^2 - H^2 - x_1^2)^{1/2} \tag{2}$$

If these coordinates are now required in along and across track coordinates, the axes have to be rotated through the instantaneous yaw angle

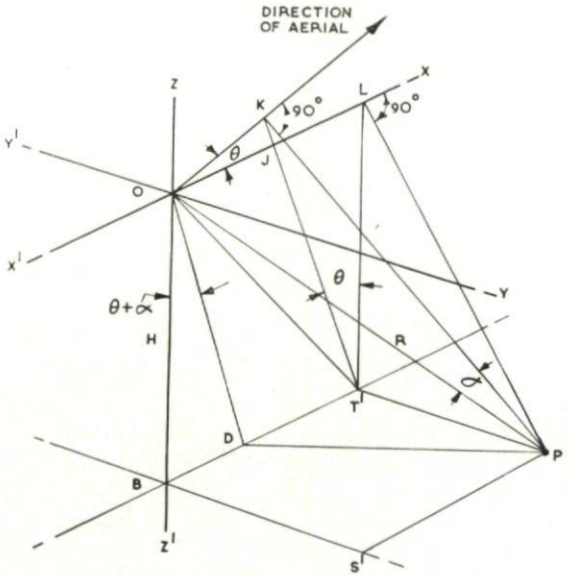


FIG. 12. *Geometry relative to heading of aircraft.*

(includes drift) in a horizontal plane, so that the new coordinates (x_2, y_2, z_2) are given by

$$x_2 = x_1 \cos \psi + y_1 \sin \psi$$

$$y_2 = y_1 \cos \psi - x_1 \sin \psi$$

$$\text{and } z_2 = z_1 = H$$

3.9. *Effect on Picture of Yaw and Roll*

Rolling of the aircraft upsets the balance of signal strength versus range set up by means of aerial shaping in the vertical plane and swept gain, so that the signal strength from equal effective reflecting areas across the picture is no longer constant. The effect is to introduce variations of intensity across the picture and in a steep angle of bank it obviously results in complete loss of the picture. A few degrees of roll are usually quite tolerable, so that no steps need be taken to minimise the effect, but where large angles of roll occur regularly it is necessary to introduce roll stabilisation.

Yaw produces the most marked effect and when varying in severe buffeting conditions it can cause undesirable distortion of the picture. A sudden

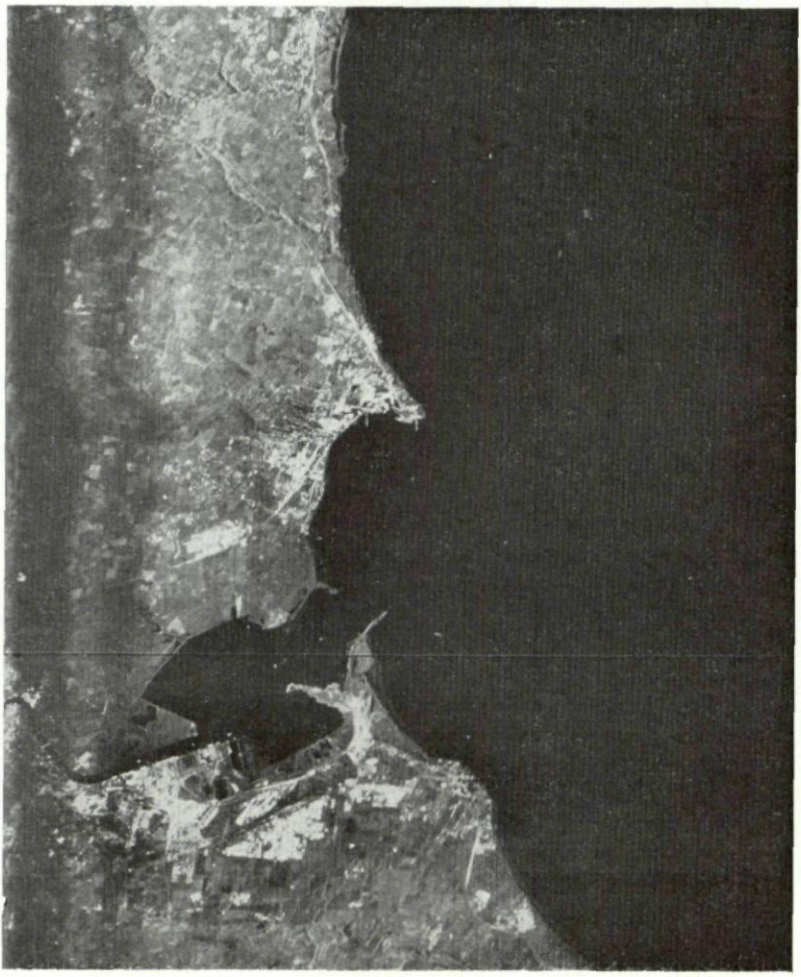


FIG. 13. *Hartlepool and Middlesbrough.*

swing of the aircraft over an angle of a few beamwidths can cause a severe reduction or even complete loss of information in a particular direction. An acceptable criterion of performance is obviously that a point on the ground should intercept the aerial beam for a time sufficient for a reasonable number of signals to be returned. On the other hand a point on the ground may be illuminated by a swinging aerial beam for a much longer period than it would with a fixed beam, and during this time the signal from it may fluctuate through more than one maximum. Acceptable effects of yaw are that the effective beamwidth should not be greater than twice the true aerial beamwidth and that the signal should not fluctuate by more that 3 db when in the region where it is a maximum. The last requirement is

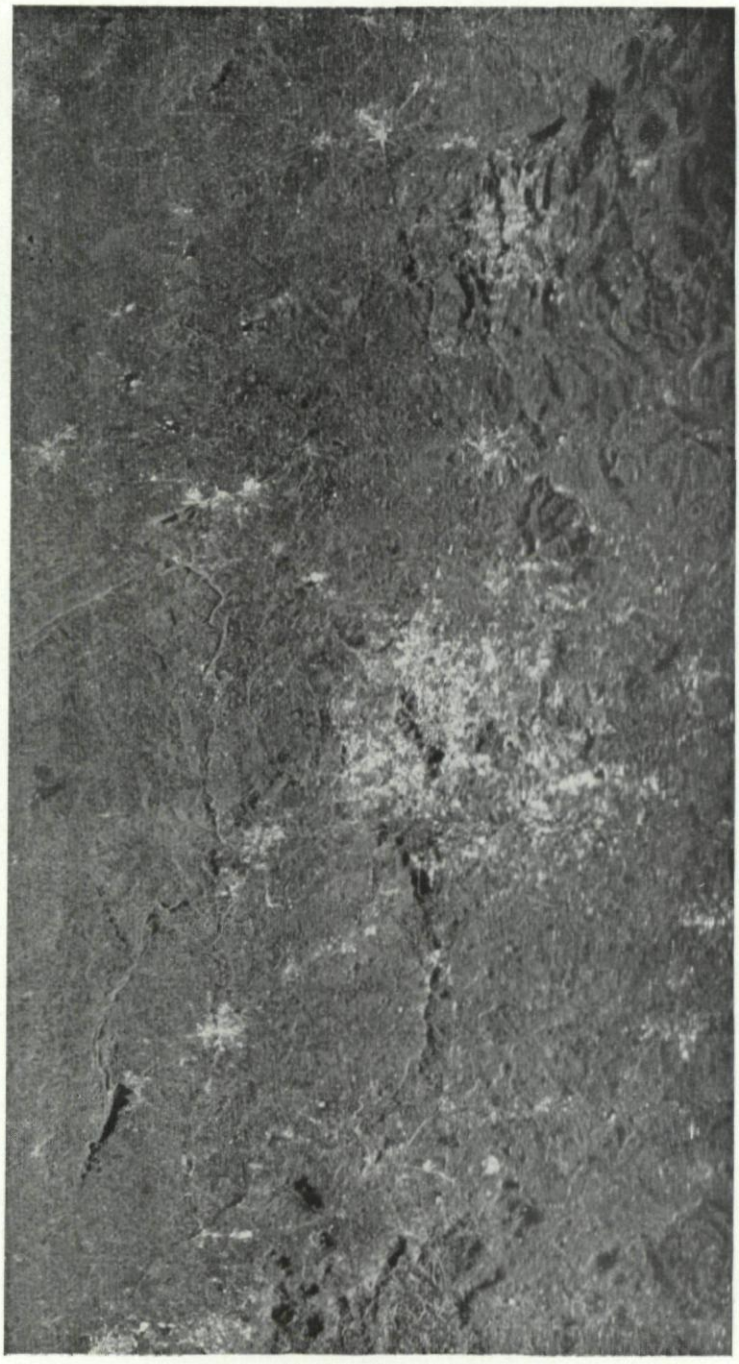


FIG. 14. *The West Midlands.*

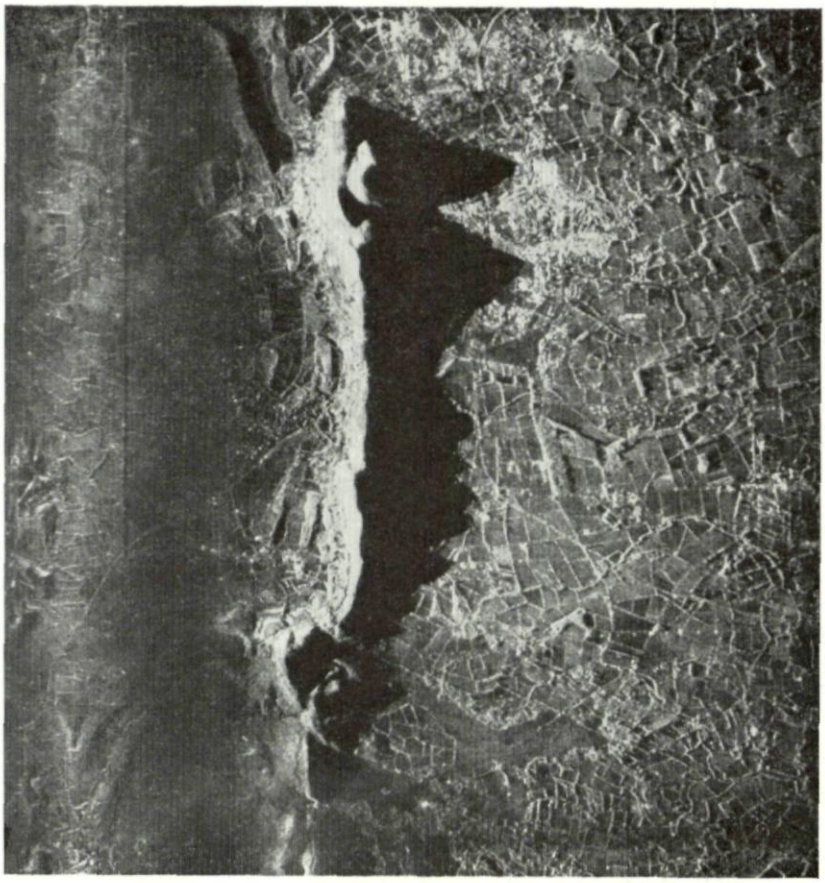


FIG. 15. *Silhouette of the Malvern Hills by sideways looking radar.*

somewhat subjective in character, and sets a limit to high frequency fluctuations, which in any case tend to be small. A tolerable amplitude for high frequency motion is about one third of a beamwidth.

Figures 13 and 14 show the excellent performance obtainable using equipments which embody the techniques discussed in the preceding pages. Figure 15 is of local interest to us at R.R.E. in that it shows the shadow cast by the Malvern Hills when illuminated by the sideways looking radar of a low flying aircraft.

DISCUSSION

J. SALOMON: Pouvez-vous faire une comparaison entre les radars á vision laterale du type non-coherent et du type coherent?

J. HIGHCOCK: In the first place the coherent radar is best used at long range where the target is in the beam for a long time and coherent integration can be used most effectively to narrow the effective aerial beam width.

For short ranges I don't think there is any particular advantage in using coherent radar. The type of radar I have been talking about this morning is most useful up to ranges of about 20 miles. Beyond that the azimuthal resolution falls off very nearly linearly with distance and becomes quite poor. This is the region where the coherent radar with its so-called synthetic aperture really pays off and there is no doubt that at say 50 miles the coherent radar wins hands down. The effective beam width of a synthetic aperture at this distance can be quite small, the linear azimuthal resolution can be theoretically somewhere about 5 or 10 feet, in practice it is probably about 50 feet. The non-coherent radar will be hard put to it to do better than about 500 feet. I think that's probably the most significant comparison I could make.

CHAPTER 12

AIRBORNE EARLY WARNING RADAR: SOME BASIC CONSIDERATIONS

SOL MATT
General Electric Company, Utica, New York, U.S.A.

1. THE BEGINNINGS

In 1935 the British Air Ministry authorized the installation of five radar stations on the east coast of England. In 1937, 15 more were added to assure radar surveillance of the whole east and southeast coast so that by the outbreak of World War II there was a continuous 24-hour radar check of the primary air and sea approaches to Britain.

FIG. 1. *The TBM Avenger, the first aircraft to carry a high-power search radar.*

So effective was this network that it has been generally credited with making it possible for the outnumbered RAF to seek out and destroy the attacking bombers, which in time switched to night bombing to reduce their losses. This tactic induced the British to mount radar equipment in their aircraft so that they could better train their guns on an adversary they couldn't see. Here was the first airborne radar.

After the war, it was revealed in military critiques that if the bombers had assembled at low altitude and penetrated at these altitudes, the RAF would have been far less effective than it was. The realization of this need—actually as early as 1943—to include low-altitude coverage in the detection of enemy aircraft was the birth of the Airborne Early Warning System.

The earliest application of this new technology came near the end of World War II with the TBM Avenger (Fig. 1) being modified to carry a

high-power search radar. This set was intended to extend the line-of-sight coverage against attacking aircraft in the Pacific. The Avenger system, was conceived as a remote, raised antenna-receiver-transmitter group. Data were transmitted from the low altitude aircraft to the surface by transmitting raw video and the antenna-pointing angle.

In 1944-45, when the Kamikaze type attacks were begun, the U.S. Navy recognized the need to use remote airborne picket stations at long distances from the beachhead for early warning and for fighter control. A crash program was undertaken to modify the B-17G as an airborne Combat Information Center. The result was the PB-1W, shown in Figure 2.

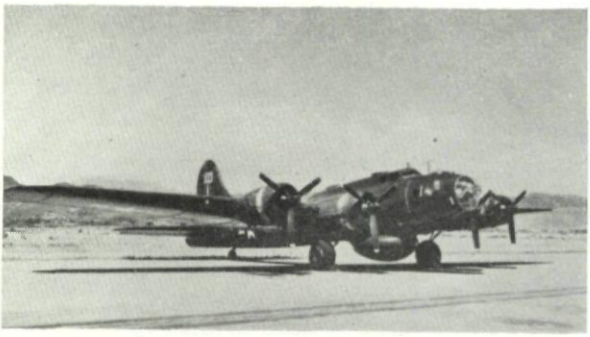


FIG. 2. *The PB-1W, a modified B-17G used as an airborne Combat Information Center in the closing days of World War II.*

Both of these systems were developed for detection of low-flying aircraft, the TBM Avenger as part of the fleet Combat Information Center system and the PB-1W as an autonomous control center.

From these beginnings, the AEW system has earned a reputation for great strategic and tactical importance. Besides extending the range for low-altitude threat detection, the AEW aircraft allows for remote positioning of a radar station over an area which is geographically inaccessible or one which is in or dangerously close to enemy territory. With its mobility, the airborne system can be deployed rapidly to meet new tactical situations, evade hostile forces, or cover large geographic areas effectively. The latter application is illustrated by the long-range barrier patrols employed for U.S. Air Defense.

2. THE MODERN AEW SYSTEM

AEW systems development has been strongly supported by the U.S. military services. The concept of the carrier-based AEW aircraft for early warning and of land-based, mobile, autonomous Combat Information Centers was further developed following World War II. Currently, the two types of systems have been combined in the Navy's E-2A Hawkeye. (Fig. 3). The APS-96 radar and detection system provides the prime sensor. The use of digital computers and automated data links provides the capability to

detect, evaluate, and report a large number of tracks. At the same time, several interceptors can be controlled from this carrier-based aircraft.

To understand the complexity of the typical modern AEW system, consider some of the functions and capability of the E-2A aircraft. As a carrier-based, airborne, tactical data system, it is the first such system built as an

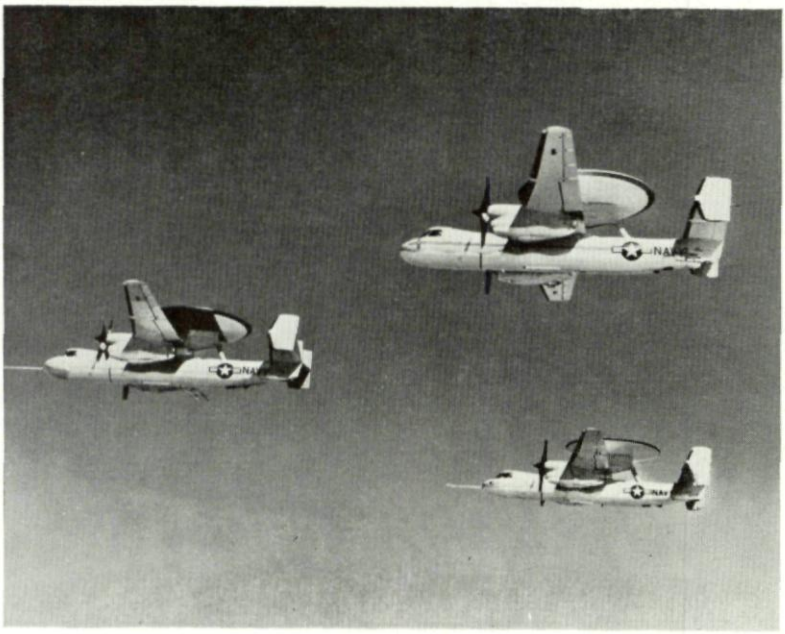


FIG. 3. *The latest in AEW aircraft, the U.S. Navy's E-2A Hawkeye.*

integrated package—both aircraft and electronic systems. It is indeed a far cry from the days of World War II when a search radar was simply stowed on board an existing aircraft and modified and adjusted until it could do the job. This E-2A effort is typical of the modern AEW aircraft which is the end product of a multitude of systems trade-offs and integration.

The E-2A is the first Navy AEW plane able to operate at high altitudes (30,000 feet and higher). Besides detecting enemy aircraft, it also gives direction for waging air battle.

The main subsystems include: radar rotodome antenna, search and height-finding radar, IFF (identification friend or foe), automatic detector, data processor (computer indicator), multipurpose communications, high-frequency data communication link, USC-2 UHF data link, and navigation subsystem (inertial, Doppler radar, and air-data computer).

The E-2A automatically or semiautomatically performs target detection, acquisition, tracking, identification, threat evaluation, and weapon assignment and intercept control. Target position, velocity, and identity are continuously displayed as symbols on a cathode ray tube. On demand,

supplementary data for a selected target are displayed on auxiliary status board readouts. Track and decision data are transmitted to surface ships for action.

The AN/APS-96 search and height-finding radar uses the pulse-compression technique—a narrow pulse is expanded in transmission and recompressed in reception. The net effect is equivalent to multiplying transmitted power by compression ratio.

Very short effective pulse length permits height finding by time-difference method. Direct target return and sea reflections are received as separate

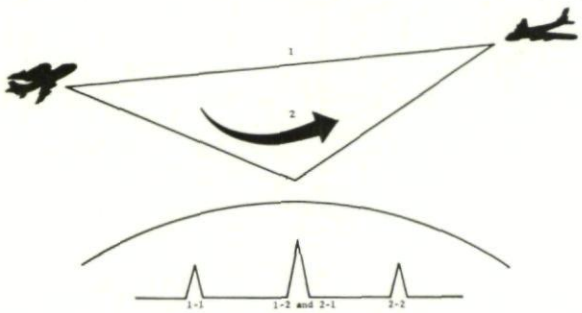


FIG. 4. *Time-difference height finding is based on the evaluation of successive radar echoes from the target.*

echoes. Together with known information, target altitude is computed. Airborne-target-indication circuits cancel unwanted clutter, especially sea-surface return. Noise figure is minimized by a parametric amplifier.

The CP-413/ASA-27 receives raw radar video, IFF video, and other navigational and antenna-bearing data, detects targets, and determines target location and height.

Target height is determined by evaluating successive radar echoes from the target during the same radar interpulse period. Normally, as illustrated in Fig. 4, three echoes may be expected from one target per radar interrogation. The first echo (1-1) is direct reflection from the target; the second (1–2 and 2–1) is target reflection that has bounced off the water on the way back to the plane, or vice-versa; the third echo (2–2) bounces off the water both ways. These three echoes arrive back at the plane at successively later intervals. Target height is derived from height of transmitting plane, range of target, and the spacing between echoes.

Beam-splitting techniques provide azimuthal accuracies over a wide dynamic range of radar and IFF return.

3. SYSTEM SYNTHESIS OF AN AEW SYSTEM

A product as efficient as the E-2A system is the result of painstaking synthesis of a multitude of system considerations having to do with both radar

and aircraft effects. Aircraft speed and altitude, for example, have a profound influence on the selection of radar parameters. The final configuration of the aircraft itself is in turn a function of, among other things, the degree of command and control that must be perfomed by the AEW system.

In a simple system where unprocessed or semi-processed data are transmitted to another location for processing, space sufficient for the radar, encoding or processing equipment, data communications, and limited monitoring facilities only need be included. When the tactical mission requires a higher level of evaluation and decision, space must be available for computers, status displays and supplementary communication. The relatively recent development of digital computers and microcircuits allows a high degree of equipment sophistication in a small volume and with low weight.

Time-on-station and radius of station or barrier length requires that considerable fuel be carried together with long-range communication equipment. If the cycle time becomes large, provisions for relief crew members as well as bunk and galley facilities must be provided. To keep the extensive array of electronic equipment operating reliably, space must also be provided for performance monitoring equipment, in-flight maintenance facilities, and spare parts.

The type of base from which the aircraft will operate (e.g. short-field or carrier requirements) also influences the aircraft design and further imposes design constraints on the detection system. In general, system effectiveness is maximized as time-on-station and radar range increase, but a point of diminishing return is soon reached. In order to detect low flying vehicles, a high station altitude is required. Hence, it is highly desirable to match the radar horizon with the maximum detection range. This is not the case for high-flying vehicles. Propagation studies have shown that there is a wide divergence of radar ranges past the horizon because of the variation in prevailing propagation conditions. Severe attenuation of target return thus results beyond the horizon.

Although the AEW antenna site is at high altitude, the apparent line-of-sight propagation path is seldom realized. The structure of the atmosphere is one of the main causes of deep fading with the horizon. A horizontally homogeneous atmosphere is rarely, if ever, realized because of synoptic meteorological conditions that are perpetually present. Stratifications caused by this synoptic meteorological pattern give rise to field-strength fading within the horizon by defocusing the lobe pattern of the transmitter along a given path. Under proper circumstances, a "ducting" condition exists, and rays tend to be "trapped" or guided within the duct. Figure 5 shows the resulting shadow zones.

For example, typical trade wind, one-way propagation ranges have been observed to vary from 500 to 1200 miles within an elevated duct, compared to less than 400 miles observed with the same equipment outside the duct. Depending on the extent of the anomaly, the effect on radar is similar to the visual effect of looking through a pane of glass, except that the atmosphere layer is not a plane, but a somewhat spherical surface parallel to the earth's surface.

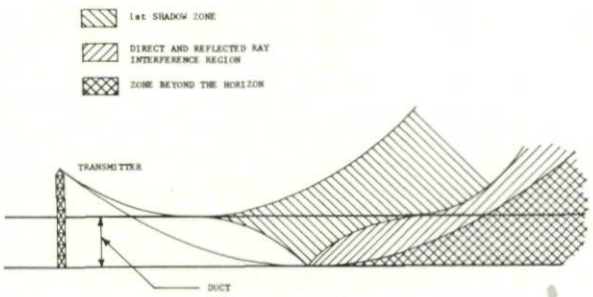


FIG. 5. *Regions of fading from different effects of radio rays.*

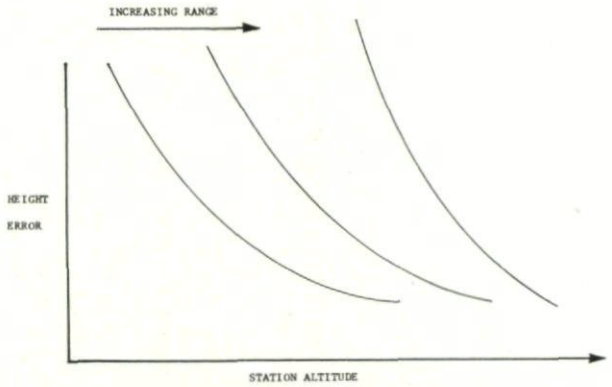


FIG. 6. *As station altitude increases, height error goes down, and ranges increase.*

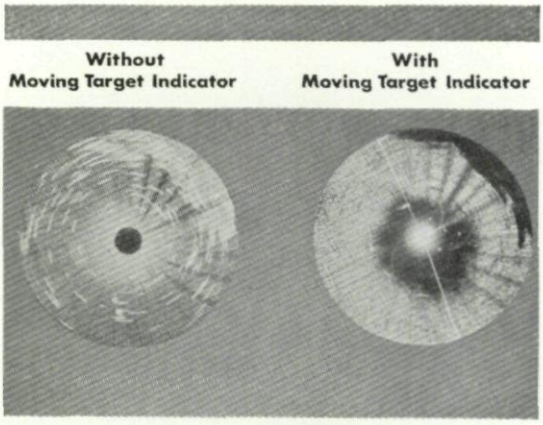


FIG. 7. *A radar scope with and without a Moving Target Indicator.*

As station altitude goes up, height error goes down and range increases. Given these circumstances, high altitude operation is preferable. For one thing, as platform altitude goes up, there is less likelihood of direct and ground-bounced interference rays. For another, when the aircraft is low it is far more likely to be in an atmospheric duct. Finally, when time-difference height finding (see Fig. 4) is used, the accuracy increases as the altitude increases. Figure 6 summarizes the advantages of high altitude operation. But, an increase in altitude brings other kinds of problems. Crew and equipment compartment air conditioning and pressurization become necessary and either peak transmitter powers are limited or radome pressurization becomes necessary.

A major limitation of AEW detection is that the price of seeing low altitude targets is extensive areas of ground and sea clutter. The magnitude and range extent of this clutter require that a high performance Moving Target Indicator be available. Figure 7 shows a radar scope before and after the use of a Moving Target Indicator (MTI). The scope photo at the left is cluttered with false targets. The photo at the right shows how an MTI rejects clutter and leaves only actual targets to be tracked.

4. THE CLUTTER PROBLEM

The major difference between an AEW search radar and a tracking or fire control radar is that the AEW radar searches an extended area and hence, even if target density is low, the number of targets will be large. A common error in system thinking is to consider only the raid model.

Friendly forces and stray traffic must be kept under constant track in order to prevent evaluation and re-evaluation as possibly hostile tracks. This usually dictates scanning the full 360° over the altitude region of interest. Even with certain geographic directions being excluded from considerations in given tactical conditions, the aircraft barrier pattern and the need for coverage during turns dictates that the antenna system have full 360° capability.

The need for high confidence, continuous coverage of a large number of targets requires an AEW system to be non-range redundant. This means the maximum pulse-repetition frequency is determined by the required surveillance range.

The time allowed for searching the coverage volume depends on many factors including the penetration that can be allowed by a hostile target, optimizing, cumulative probability of detection, obtaining accurate velocity estimates, preventing track dropping or "misassociation," updating any data to be reported to other control centers, detecting target maneuvers, and obtaining high probability of successful weapon control.

One method of determining targets from clutter is based on the movement of the targets of interest in relationship to the clutter. In an airborne system, compensation for the motion of the platform must be made. The necessity for unambiguous range on a large number of targets generally precludes the use of continuous wave or high PRF pulse Doppler systems.

Delay-time filters, banks of comb filters, and on-range gating and discrete filters are the common methods of moving target indication. Because of the narrow pulse widths usually employed and the long detection-range requirement, the latter two methods are less satisfactory, requiring a large bulk of hardware plus the requirements for large dynamic range and isolation. Multiple delay-line filters can provide high clutter rejection and the use of pulse-repetition frequency stagger can minimize the effect of blind

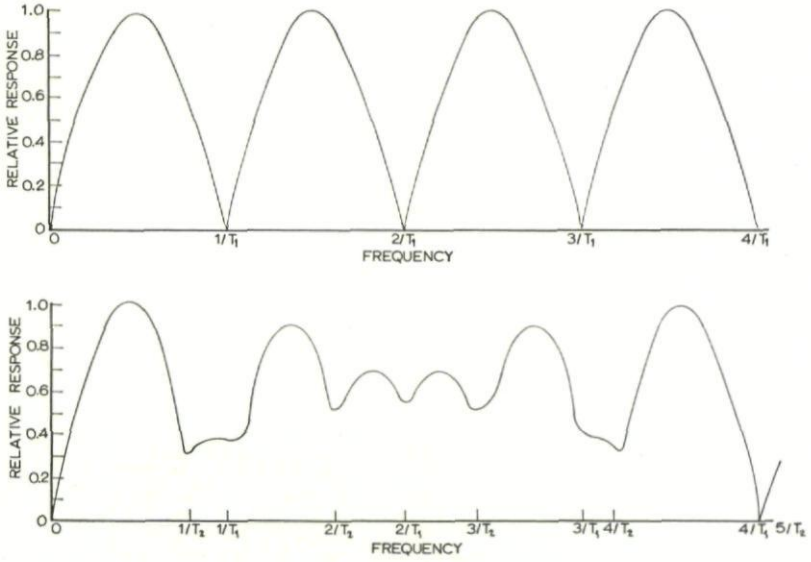


FIG. 8. *Frequency response characteristic for* $f_r = 1/T_i$ *is shown (at top) and composite frequency response with* $T_1/T_2 = 4/5$ *(at bottom).*

speeds if redundant velocity results from maximum range and target speed requirements. The frequency-response characteristic for $f_r = 1/T_i$ is shown in the upper portion of Fig. 8 and the composite frequency response with $T_1/T_2 = 4/5$ in the lower portion.

The ability of an MTI to cancel clutter is determined by the ratio of the total spectrum to the pulse-repetition frequency and the filter shape (Fig. 9). Higher order filters yield sharp skirts but are limited by the stop-band rejection.

The internal motion of clutter is due to the distribution of velocity of the scatterers. Land clutter, fortunately, has a narrow spectrum which allows its high energy content to be filtered. As shown in Fig. 10, sea clutter is much broader and has less energy than land clutter. Since Doppler frequency is proportional to velocity, the spectrum of the clutter scatterers is proportional to frequency.

If the antenna scans continuously, the amplitude modulation of the antenna pattern gives rise to a scanning spectrum. The width is inversely

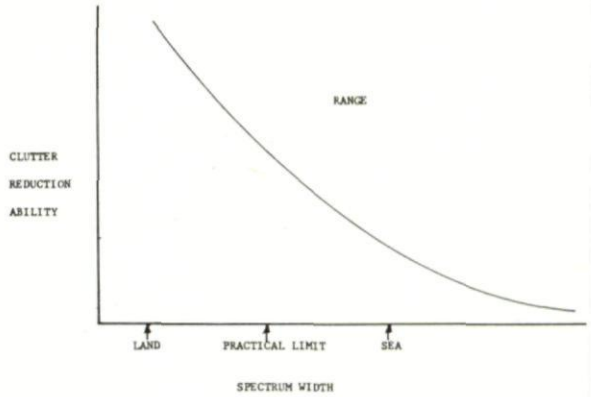


FIG. 9. *The ratio of the total spectrum to the pulse-repetition frequency and the filter shape determines the MTI's ability to cancel clutter.*

proportional to antenna beamwidth, which means that the spectrum width is proportional to frequency and antenna aperture. This relationship is represented by the following formula:

$$\sigma \text{ internal motion} = k_0 F_0$$

$$\sigma \text{ scan} = k_1 A_h F_0$$

$$\sigma \text{ platform motion} = \frac{k_2 V_g \sin \alpha}{A_h}$$

A_h = Horizontal Operation

F_0 = Radar Frequency

V_g = Ground speed

α = Antenna boresight in reference to ground track

Although the average motion of the clutter in the antenna beam can be removed, the differential Doppler of the scatterer off the antenna boresight

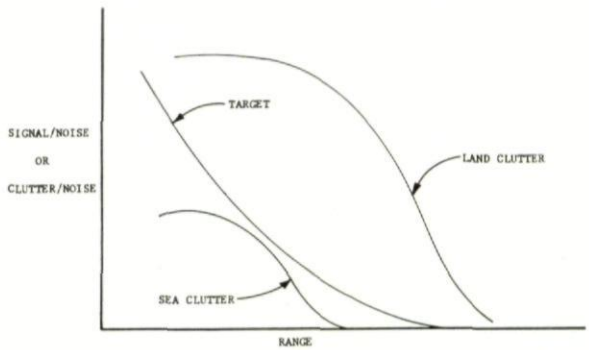


FIG. 10. *Sea clutter is broader and has less energy than land clutter.*

gives rise to a platform motion spectrum. This spectrum is proportional to the component of aircraft ground speed perpendicular to antenna boresight, frequency, and horizontal antenna beamwidth. Since beamwidth is inversely proportional to aperture and frequency, the platform motion spectrum is invariant with frequency, being proportional to ground speed, sine of the angle of boresight in respect to ground track, and inversely proportional to antenna aperture.

It is possible to compensate for scanning and platform motion since these effects are systematic with the effective angle of the scatterer in a resolution cell off antenna boresight.

As Fig. 11 shows, the ability to reject these effects decreases with frequency, the factor of ground speed perpendicular to boresight, antenna scan rate,

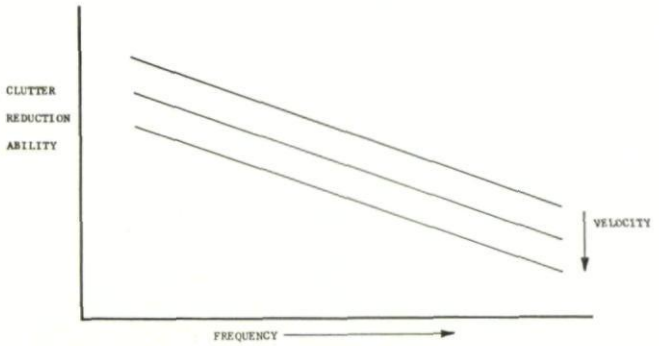


FIG. 11. *The ability to reject scanning and platform motion decreases with frequency antenna scan rate, and interpulse period.*

and interpulse period. Since the scanning spectrum is proportional to antenna aperture and the platform motion spectrum is inversely proportional to aperture, we find a large region that is relatively insensitive to antenna aperture.

Electronically scanned arrays at first glance appear attractive since the scanning spectrum can be eliminated; however, the pattern distortion that occurs when the beam is pointed off center will not allow for effective platform motion compensation.

Another limitation results from system noise which arises from timing and frequency instabilities, amplifier distortion, component imperfections, antenna pattern distortion, and delay line spurious response. The effect of these hardware imperfections is generally proportional to frequency and interpulse period and inversely proportional to pulse length.

Still another major limitation results from antenna sidelobes. Since the sidelobes point in all directions, particularly when the antenna is mounted on the aircraft, the resultant spectrum is weighted by the sidelobe response and extends from the negative value of the aircraft ground speed to the positive value. Even if target speeds are high compared to the AEW aircraft, the

radial velocity distribution extends to zero and, as shown in Fig. 12, is quite significant.

This means that the clutter filter will see sidelobes unless it rejects targets that are not closing or opening at speeds greater than own ship's velocity. Except in very specialized applications, these conditions do not exist for a search radar.

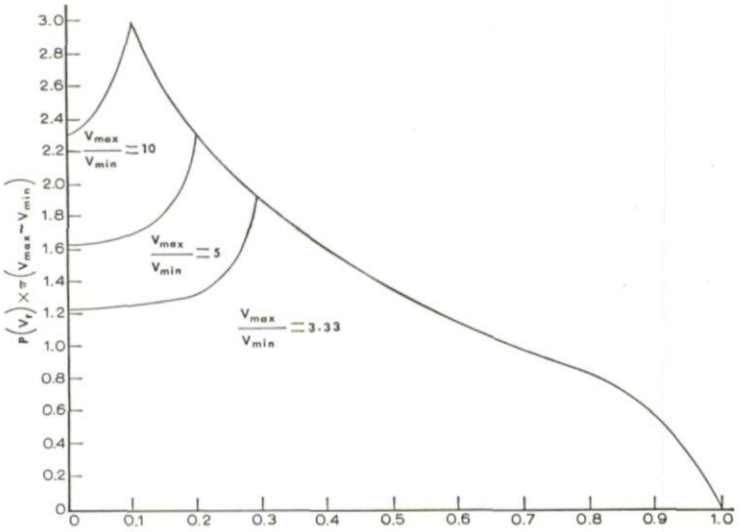


FIG. 12. *Probability-density functions of relative velocity, assuming target velocities distributed uniformly from* v_{min} *to* v_{max}.

This also means that the antenna must be constructed with low sidelobe levels, and the installation on the aircraft must be chosen to minimize degradation of the antenna sidelobes. Since 360° coverage is generally required, the antenna beam should be displaced as far as possible from the aircraft structure. Realization of low sidelobes requires very tight antenna tolerances. For a well designed and constructed antenna, as frequency increases, the electrical error will increase proportionally since our ability to maintain tolerance is usually on a finite-dimension basis.

As far as clutter rejection is concerned, it is desirable to use low radar frequencies; however, antenna beamwidth is inversely proportional to frequency. A narrow beam contains less clutter, but the degradation of performance is a higher order of frequency than the first power.

Narrow beams are desirable from the standpoint of angular resolution and accuracy. At first glance the radar range equation appears to indicate that an increase of frequency is desirable since gain increases for constant antenna aperture; however, since a search radar requires a given altitude coverage, the vertical beamwidth cannot be reduced. This causes any gain increase due to horizontal beam narrowing to be cancelled as the vertical aperture is

reduced to maintain constant beamwidth. If a fully coherent radar detection system is available, the detection range is independent of frequency.

For more practical systems where the range is proportional to the square root of the number of hits, we find that noise-limited detection range decreases as the eighth root of frequency.

The choise of frequency is a compromise of angular resolution versus detection and anticlutter performance. Large horizontal apertures can give better resolution and noise-limited detection range. If short pulses are used to obtain range resolution, then effective angular resolution is obtained except when targets are range coincident.

5. CONCLUSION

These then are some of the basic considerations in designing airborne early warning systems. Although the point has not been stated explicitly so far, one conclusion is inescapable. Any military system as sensitive and as complex as an AEW system cannot be an afterthought to be installed on an existing aircraft. It must be carefully designed into an aircraft, and in all instances both the radar and the aircraft designer must coordinate their efforts from the beginning, within the limitations imposed by hardware, state-of-the-art, and propagation phenomena. Only in this way will the total AEW system perform its mission efficiently and reliably.

6. DISCUSSION

———— You made a comparison between the spectrum widths of land clutter and sea clutter. Could you tell us what the sea state is for that particular value you quoted. Alternatively, how the land clutter compared with sea clutter at the various sea states.

S. MATT: The slide showed a general curve and did not indicate any specific sea state. However, we have found that in most cases, land clutter exceeded all relevant sea states which we have measured up to sea state 5. There are a few cases where land clutter is actually less than sea clutter. I would guess less than 20%. We have always found that the land clutter has a much narrower spectrum than has the sea clutter from the same range. Here you must be careful because the range from your station is important.

N. J. SMITH: In connection with the desirability of high altitude operation, have you any views or experience of the variation of signal/clutter with altitude? Theory, with some experimental support suggests that for a conventional pulse or pulse compression radar the average signal/clutter ratio will fall off for a given range at least proportionately to height. This is not the case for C.W. radars where an improvement with increased height is predicted and observed.

S. MATT: Our experience in flying over water gives some improvement with height. However over land the matter is complex. The clutter level can vary by 30 db with change in terrain.

N. J. SMITH: In the comments on land and sea returns were you primarily dealing with small glancing angles?

S. MATT: The clutter was at low angles corresponding to a 40 mile range.

R. W. WILMER: I wonder if you could, without getting into security difficulties, tell us very broadly what success you had with this method of height determination—perhaps I should be specific and say, over the average rate of sea clutter roughly what is the chance of being able to get good height? Can you get height at all over land?

S. MATT: Over sea, we have had excellent results of being able to determine the altitude—over land, our present system which was designed for over sea operation is not doing so well. I can only guess that it is about 20%, though the equipment is being modified to help the situation.

CHAPTER 13

AIRBORNE DOPPLER NAVIGATION TECHNIQUES

TREVOR GRAY
Decca Radar Ltd., Isle of Wight Division, England

1. INTRODUCTION

It is now between 10 and 15 years since it was shown that airborne Doppler radar navigation systems were a practical proposition. Since that time many specific systems have been designed and put into operation and many companies and organisations have published information on specific systems seeking to extol the virtues of one or other type of system.

Sufficient time has now elapsed for an objective view of the situation to be taken. There can be little doubt that in the latter part of the 1950's the possibilities of Doppler navigation were oversold. As a result of this, when practical systems came into operation, an inevitable reaction against Doppler systems occurred. However, there is now sufficient evidence from a considerable amount of flying for an assessment of the capabilities of Doppler to be made. It is clear that a Doppler system requires, for its most advantageous use, a combination of a good heading reference system and a good computer. These points were not always fully realised by potential users in the early stages of Doppler use and will be referred to later in this paper.

The purpose of this paper is not to provide a significant amount of new material but rather to collate information on the techniques which have been known for some time, and to make an effort to indicate the advantages and disadvantages of each technique for specific purposes. The succeeding section of the paper will be sub-divided in the following way.

A brief description will be given of possible system configurations and this will be followed by a description of the various transmission techniques available to the designer, followed by discussion of other variable parameters. A section will also be included on the methods of computing the information provided by the basic Doppler system. This will be followed by a section indicating the effect of recently developed components and techniques on Doppler design.

2. SYSTEM CONFIGURATION

Any airborne Doppler navigation system must comprise an antenna system, a transmitting and receiving system, a frequency measuring system and some form of readout or display. It may be that a computer will also be included but this is not inevitable. It is proposed to deal first with the combinations available to the designer in respect of the system in its broadest terms.

Before, however, describing such systems it is necessary to define the information which the equipment is required to provide. The Doppler navigation system is capable of providing information on the velocity of the vehicle to which it is fitted in relation to the terrain over which it is flying. The basic system can provide information in a co-ordinate system related only to the vehicle Thus it is possible to provide vehicle velocity in terms of velocity along and across the vehicle heading, together with vertical velocity. These are merely components of vehicle velocity in polar and cartesian co-ordinates with respect to vehicle heading. It should be noted that it is not possible to provide vehicle velocity information in terms of co-ordinates related to the earth unless an independent heading reference is provided.

One broad sub-division of system variables concerns the form of the antenna. It is obvious that any single beam of energy directed towards the ground is capable of providing only the component of vehicle velocity along the direction of that beam. Therefore, in order to define completely the vehicle velocity in space, a minimum of three beams is required These provide horizontal velocity in two orthogonal directions together with vertical velocity. It is, of course, equally possible to use a single beam capable of being changed in direction with time. This is, however, a trivial variant of the multi-beam conception. A discussion on possible beam configurations has been published by Fried.[1]

It is now proposed to give some indication of the choice of systems available to the designer of Doppler equipment. There are several areas in which a choice has to be made which fundamentally affects the design. These are as follows:

1. There is a choice of transmission characteristics between:
   (a) Unmodulated continuous wave,
   (b) Frequency modulated continuous wave,
   (c) Coherent pulse modulation,
   (d) Non-coherent (or self coherent) pulse modulation.

2. The receiver,
   (a) This can be either a superheterodyne or
   (b) A homodyne. This term is used to indicate a superheterodyne with zero intermediate frequency.

3. The antenna stabilisation may be broadly classified as
   (a) Drift stabilised, or
   (b) Fixed in azimuth.
   Both drift stabilised and azimuth fixed antennas may be stabilised in pitch and/or roll.

4. The antenna beam shape must be defined in general. Beams may be either:
   (a) Pencil beams, or
   (b) Conical beams providing hyperbolic interceptions on the ground.
   Number of beams and configuration will not be discussed here. This is fully treated by Fried.[1]

5. It has already been stated that more than one beam is required. This requirement may be realised either by
   (a) Operating one beam at a time sequentially, or
   (b) Operating all beams all the time, i.e. parallel operation.
6. The necessary isolation between the transmitter and receiver may be obtained either:
   (a) by a duplexing arrangement using a single antenna, or
   (b) by using separate antennas, this scheme sometimes being referred to as space duplexing.

There are a large number of other factors which have to be decided, but the above mentioned choices are the main ones. A simple calculation will indicate that there are 128 possible variants of the above parameters but not all of these possibilities would result in practical systems and so Figs. 1 to 5 have been prepared which indicate these possibilities which are considered to be practicable. The factors affecting the choice of the above variables will now be discussed in more detail.

3. CHOICE OF TRANSMITTING SYSTEM

It has been mentioned above that there are four known possible forms which the transmitter may take. These have been defined as

(i) Unmodulated C.W.
(ii) F.M. C.W.
(iii) Coherent pulse
(iv) Self-coherent pulse.

3.1. *The C.W. System*

In this system a continuous unmodulated signal is transmitted. The returned signal is mixed with a sample of the transmitted power and thus the Doppler signal is obtained. This system is very efficient because as there is no modulation of the transmitter, all the Doppler signal is contained in a single section of the spectrum; furthermore, since coherent detection may be used, the receiver noise is determined by the post detector bandwidth which may be made very narrow. Because the transmitted power is scattered by an extended target area the Doppler signal-power decreases at a rate proportional to the inverse square of altitude.

Extra noise, with spectral components in the Doppler band, is introduced into the system due to unwanted phase- and amplitude-modulation of the carrier leakage between the transmitting and receiving channels. This is caused by backscattering and vibration from discontinuities associated with the aircraft structure and radome, and by shock waves in the case of high-speed aircraft. This problem demands separate transmitting and receiving antennas in order to achieve the required degree of isolation. The transmitter frequency should be highly stable otherwise unwanted frequency modulation on the carrier frequency results in a loss of coherence and consequently a reduced Doppler signal-to-noise ratio.

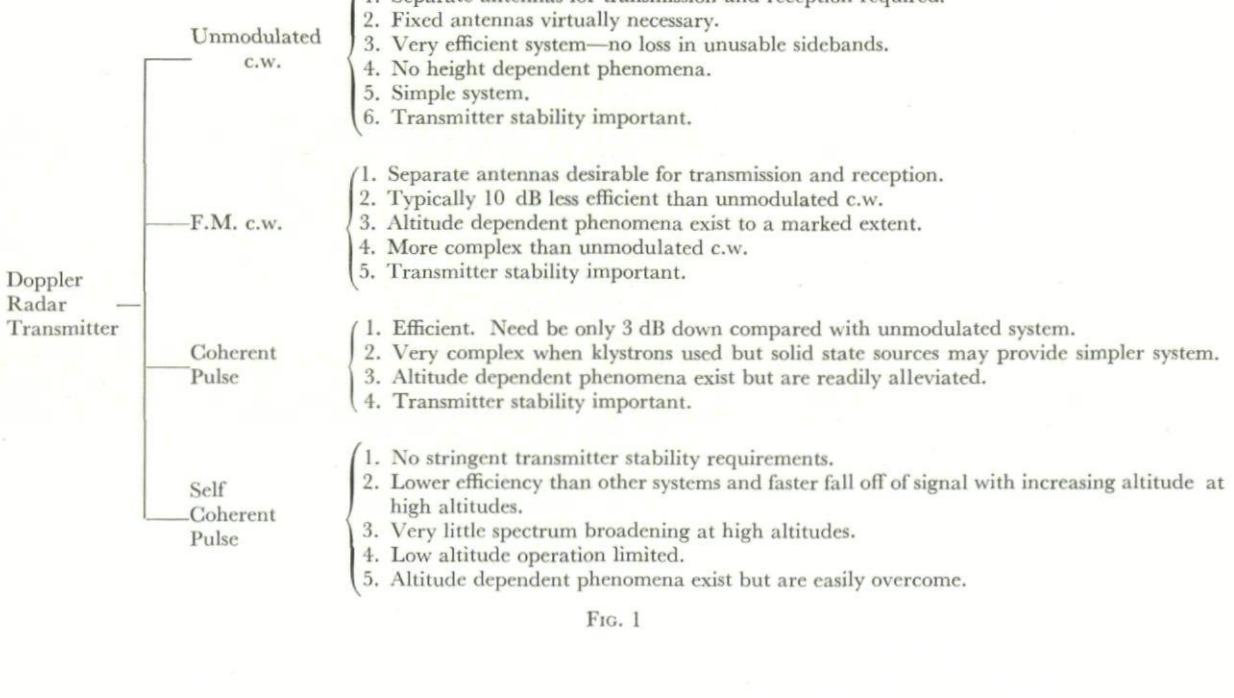

Doppler Radar Transmitter
Unmodulated c.w.
1. Separate antennas for transmission and reception required.
2. Fixed antennas virtually necessary.
3. Very efficient system—no loss in unusable sidebands.
4. No height dependent phenomena.
5. Simple system.
6. Transmitter stability important.
F.M. c.w.
1. Separate antennas desirable for transmission and reception.
2. Typically 10 dB less efficient than unmodulated c.w.
3. Altitude dependent phenomena exist to a marked extent.
4. More complex than unmodulated c.w.
5. Transmitter stability important.
Coherent Pulse
1. Efficient. Need be only 3 dB down compared with unmodulated system.
2. Very complex when klystrons used but solid state sources may provide simpler system.
3. Altitude dependent phenomena exist but are readily alleviated.
4. Transmitter stability important.
Self Coherent Pulse
1. No stringent transmitter stability requirements.
2. Lower efficiency than other systems and faster fall off of signal with increasing altitude at high altitudes.
3. Very little spectrum broadening at high altitudes.
4. Low altitude operation limited.
5. Altitude dependent phenomena exist but are easily overcome.

FIG. 1

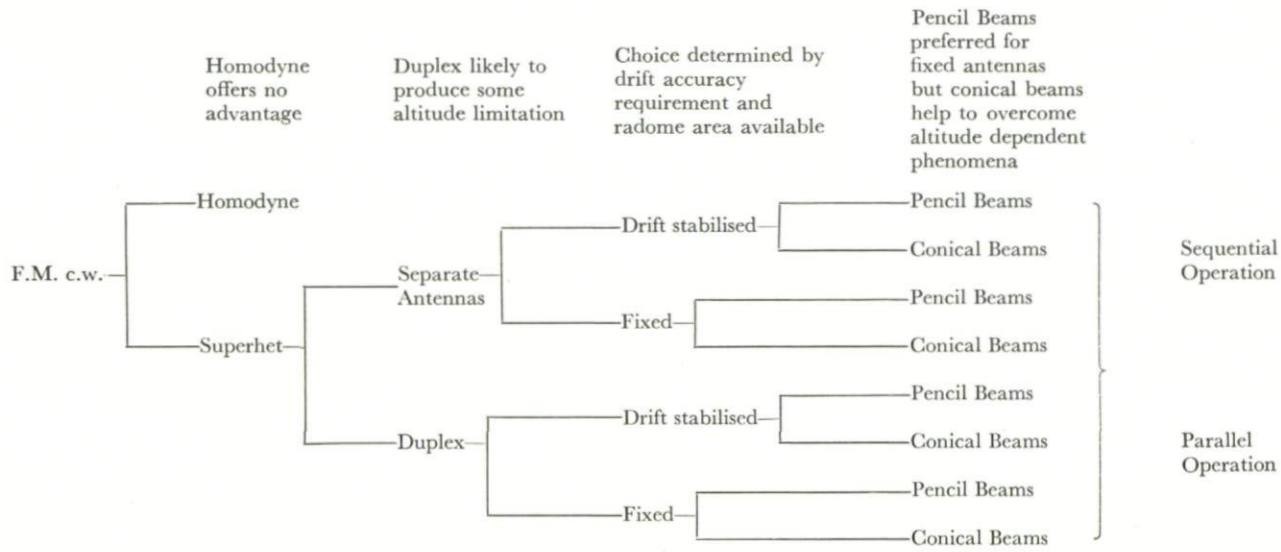

Homodyne offers no advantage
Duplex likely to produce some altitude limitation
Choice determined by drift accuracy requirement and radome area available
Pencil Beams preferred for fixed antennas but conical beams help to overcome altitude dependent phenomena
F.M. c.w.
Homodyne
Superhet
Separate Antennas
Duplex
Drift stabilised
Fixed
Drift stabilised
Fixed
Pencil Beams
Conical Beams
Pencil Beams
Conical Beams
Pencil Beams
Conical Beams
Pencil Beams
Conical Beams
Sequential Operation
Parallel Operation

FIG. 2

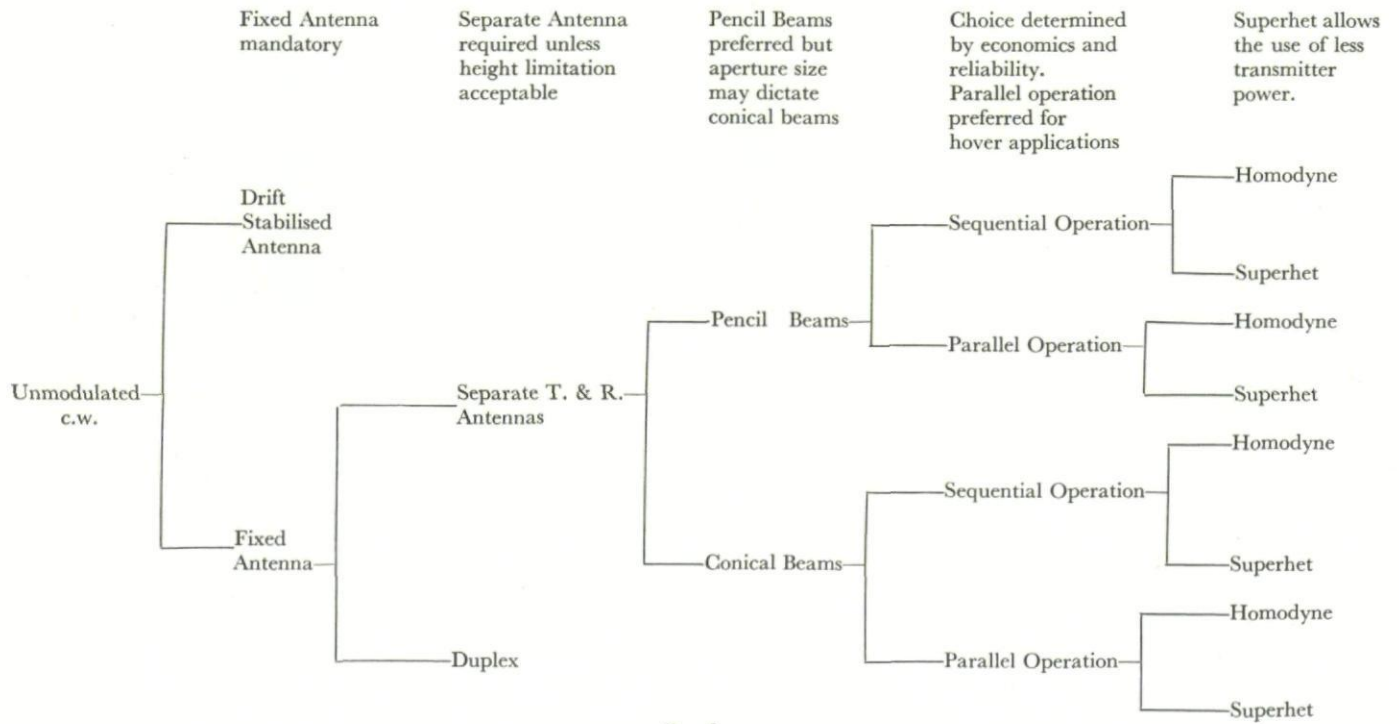

Fixed Antenna mandatory
Separate Antenna required unless height limitation acceptable
Pencil Beams preferred but aperture size may dictate conical beams
Choice determined by economics and reliability. Parallel operation preferred for hover applications
Superhet allows the use of less transmitter power.
Unmodulated c.w.
Drift Stabilised Antenna
Fixed Antenna
Separate T. & R. Antennas
Duplex
Pencil Beams
Conical Beams
Sequential Operation
Parallel Operation
Sequential Operation
Parallel Operation
Homodyne
Superhet
Homodyne
Superhet
Homodyne
Superhet
Homodyne
Superhet

FIG. 3

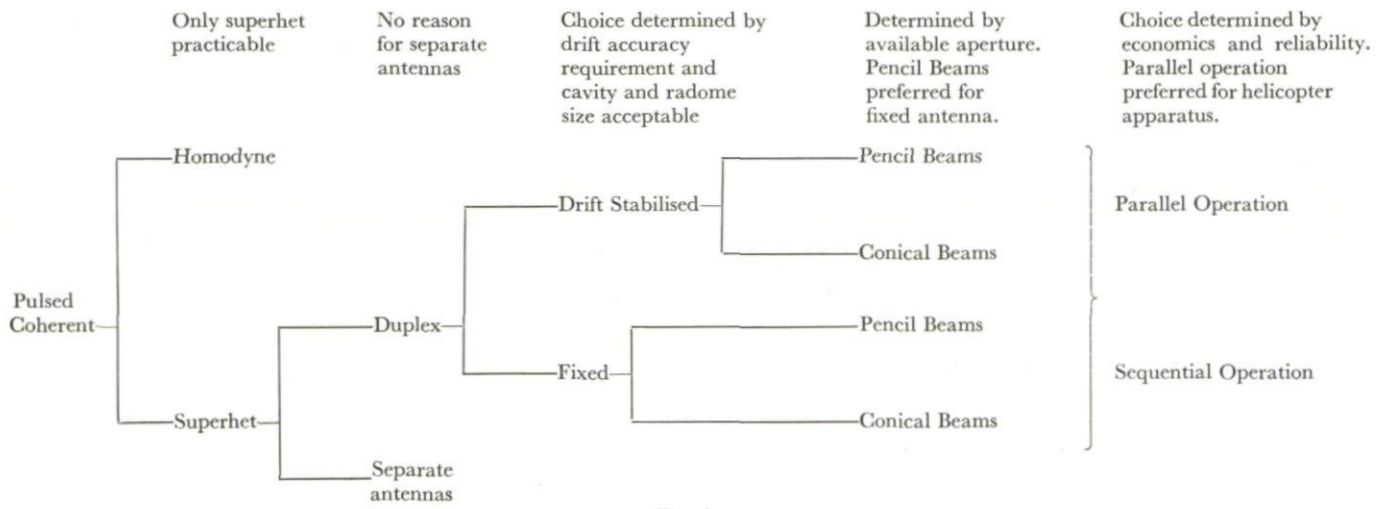

Only superhet practicable
No reason for separate antennas
Choice determined by drift accuracy requirement and cavity and radome size acceptable
Determined by available aperture. Pencil Beams preferred for fixed antenna.
Choice determined by economics and reliability. Parallel operation preferred for helicopter apparatus.
Pulsed Coherent
Homodyne
Superhet
Duplex
Separate antennas
Drift Stabilised
Fixed
Pencil Beams
Conical Beams
Pencil Beams
Conical Beams
Parallel Operation
Sequential Operation

FIG. 4

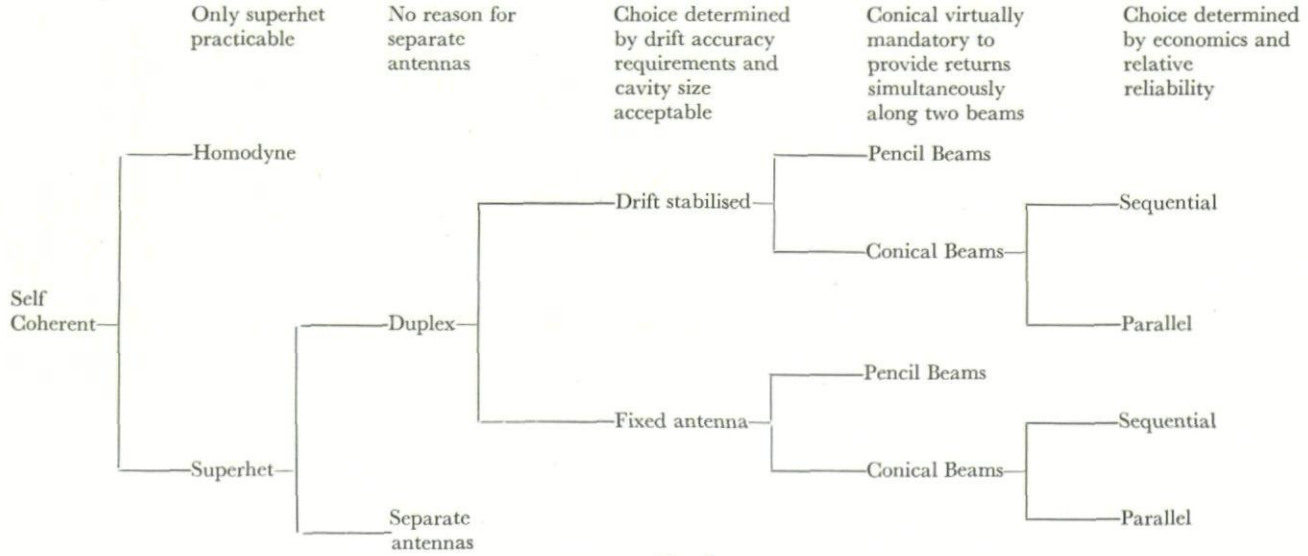

Only superhet practicable
No reason for separate antennas
Choice determined by drift accuracy requirements and cavity size acceptable
Conical virtually mandatory to provide returns simultaneously along two beams
Choice determined by economics and relative reliability
Self Coherent
Homodyne
Superhet
Duplex
Separate antennas
Drift stabilised
Fixed antenna
Pencil Beams
Conical Beams
Pencil Beams
Conical Beams
Sequential
Parallel
Sequential
Parallel

FIG. 5

3.2. *F.M./C.W. System*

The major problems encountered with c.w. systems are due to the presence of reflections from close-range targets. If, however, the transmission is frequency-modulated, a type of system emerges which offers low sensitivity to close-range targets and targets spaced at multiples of distance corresponding to successive cycles of the modulating frequency. Thus close-range targets cease to be a serious problem.

However, the Doppler energy is now no longer confined to a single part of the spectrum but instead is distributed in spectral sidebands spaced at intervals depending upon the modulating frequency. Using one of these spectral sidebands only, therefore, introduces a loss which in practical systems amounts to about 12 db compared with an ideal c.w. system.

The Doppler signal is recovered by mixing the received signal with a sample of the transmitted signal, and using a bandpass amplifier to select that spectrum which it is proposed to use. This is possible because the result of the mixing process is to produce signals at the modulating frequency and at harmonics of this with the Doppler frequency as sidebands. The modulation index is adjusted to make the power in the harmonic selected a maximum.

Coherent detection may be employed by using the appropriate harmonic of the modulating frequency as the switching signal.

Although one of the apparent advantages of frequency modulated systems is that of allowing duplex systems to be used, in many practical applications of the system, separate transmitting and receiving antennas have been found to be necessary.

3.3. *The Coherent Pulse System*

Another approach enabling the use of a single antenna is that of the coherent pulse system in which the transmitter is turned off during the period when reception is taking place. It is, however, necessary that the transmitter or some signal from which it is derived should operate continuously in order to provide a coherent signal which can be used for extracting the Doppler information from the returned signal. This system is highly efficient, being, when optimised, just 3 db down in performance compared with an ideal unmodulated c.w. system. However, the practical implementation of this system usually leads to considerable complexity. Most of this complexity results from the fact that it is necessary to prevent the output from the continuously running transmitter oscillator from entering the receiver during the receiving part of the cycle. This has been found to be extremely difficult with a continuously running oscillator at transmitter frequency, and has led to the development of systems in which the continuously running oscillator is at some frequency other than that of the transmitter, the transmitted frequency being obtained by a mixing process followed by a klystron amplifier operating only during the transmitting part of the cycle. This has resulted in transmitters of considerable complexity although it will be indicated later that forthcoming component developments may alleviate this situation.

3.4. *Self Coherent Pulse*

This system differs from the three systems already mentioned in that there is no necessity to take special measures to ensure transmitter stability. This comes about because the system is operated as a Janus system in which Doppler signal is extracted by beating together signal returns from two divergent directions originating from the same transmitted pulse. The system is fully described by Berger[2] and by Clegg and Thorne[3] and so no detailed description will be given here. The system is one of simplicity but rather lower efficiency than the other systems described, and imposes certain limitations in connection with antenna stabilisation. In particular it is necessary for the signal returns from the two Janus beams to be obtained simultaneously and thus for the best operation the system should be stabilised both in roll and in pitch. An acceptable system can be developed without roll stabilisation by using beams which are wide in the broadside direction, but even in these circumstances roll performance is worse than that experienced with the other systems described. One great advantage of this system is that it is in no way critical and very little spectrum broadening, due to coherence loss, is experienced at any altitude. Furthermore Janus operation provides a spectrum width only 70% of that obtained with a single ended system for a given antenna beam width. This factor leads to good accuracy of tracking and so this system is particularly suitable for applications requiring the highest accuracy.

4. RECEIVER

With some types of Doppler system there is a choice of receiving technique between the superheterodyne and the homodyne or zero I.F. superheterodyne. With any type of system the superheterodyne can always be used and will offer a noise factor, at the present time, of the order of 9 to 10 db. There is nothing abnormal about this type of receiver compared with conventional radar techniques and the choice of intermediate frequency normally lies between 10 and 60 Mc/s. Intermediate frequency can be obtained either with a separate local oscillator in all systems or by a side stepping technique in systems where a coherent oscillator is available. There is, however, considerable attraction about any measures which can provide simplification and this has led to some use of the homodyne receiver. This type of receiver is applicable only to the continuous wave system where injection of a sample of the transmitted frequency into the receiver produces the Doppler signal directly as an audio frequency. Thus the local oscillator and all supplies and control circuits connected with it are eliminated. The disadvantage of the homodyne receiver has been that the noise factor with existing systems is poorer than that of the superheterodyne, being of the order of 35 to 40 db. Furthermore, the noise generated in the crystal mixer is inversely proportional to frequency which means that the noise increases as the Doppler frequency reduces. This results in asymmetrical noise characteristics under poor signal conditions, thus causing an additional source of error in the tracker.

However, a considerable amount of development effort has been applied recently to the improvement of crystal mixers for this type of service and

information now becoming available suggests that homodyne receivers with performance comparable to superheterodynes may soon be a practical proposition.

The tunnel diode amplifier is another development which may well lead to improved and simplified receiving systems.

5. ANTENNA STABILISATION

Perhaps the broadest subdivision between antenna types is between drift stabilised antennas and antennas fixed in azimuth. Dealing first with drift stabilised antennas the following are the major characteristics:

(a) This technique is not suitable for use with unmodulated c.w. systems owing to the fact that the changing conditions provided by a moving aerial would react unacceptably on transmitter stability and on the degree of isolation obtainable between transmitting and receiving systems.

(b) The drift stabilised antenna undoubtedly provides the highest possible accuracy of drift measurement particularly at high angles of drift.

(c) The drift stabilised antenna enables a simple system to be used if the outputs are required in the form of ground speed and drift angle.

(d) There is a limit to the speed of response to changes of drift angle due to the relatively heavy mass of the antenna which has to be stabilised.

(e) The radome area required with a drift stabilised antenna is a function of the amount of drift coverage required and in general is considerably greater than the actual antenna aperture.

The salient features of the antenna fixed in azimuth are as follows:

(a) Installation is, in general, easier, than with a stabilised antenna.

(b) The cut out in the skin of the aircraft is smaller than with a drift stabilised antenna.

(c) High speeds of response in the system can be achieved.

(d) System accuracy particularly in drift will, in general, be lower than that obtained with the drift stabilised antenna particularly at high drift angles. There are areas where this is not of particular importance, these being fast aircraft where small drift angles only are to be expected and very low speed aircraft such as helicopters near the hover.

It has already been mentioned that both types of antenna may have pitch and/or roll stabilisation added. It is, however, unusual for pitch or roll stabilisation to be added to the azimuth fixed antenna since this arrangement produces almost all the disadvantages of the drift stabilised antenna without offering the advantages of that system. One exception to this general statement is that a fixed antenna used with a self coherent pulse Janus system would probably require pitch stabilisation for reasons already stated.

6. ANTENNA BEAM SHAPE

The two general types of beam shape used in Doppler systems are pencil beams, which have a similar beam width in both directions, and beams which are wider in the broadside direction than in the fore and aft direction. Dealing first with pencil beams, these are particularly applicable to fixed aerial systems in which high drift angles are to be expected since spectrum broadening, due to drift is minimised by this beam configuration. This type of beam is unsuitable for self coherent systems as an intolerably high degree of antenna stabilisation would be necessitated. The use of beams elongated in the broadside direction represented a considerable step forward in the Doppler art since it was found that a conical beam produced by a suitable linear array, producing as it does a hyperbolic intercept on the ground, offers the possibility of a very narrow spectrum since lines of constant Doppler frequency are also hyperbolic. Thus by using a combination of wide broadside beams which are narrow in the fore and aft direction and using azimuth stabilisation to make the beams and isodops coincident a narrow Doppler spectrum and a good degree of roll tolerance are simultaneously achieved.

This invention has been particularly important to the self-coherent type of system and is described by Berger.[2] It should be noted that the precise depression angles of the antenna beams are of paramount importance to the accuracy of the system. It is found to be exceedingly difficult to measure beam angles on ground equipment to the necessary degree of accuracy to enable the calibration constant of the resulting Doppler system to be predicted This has led to the development of antenna designs such that beam positions can be controlled solely by readily measured mechanical dimensions. It has been found that slotted linear arrays of waveguide fall into this category and it is only with this type of antenna so far that it has been possible to produce systems of the highest accuracy without the necessity for flight calibrating individual antennas. There is a further advantage of the slotted linear array in that a suitable combination of phased and anti-phased radiators provides a system where the calibration constant is independent of frequency and largely independent of ambient temperature. This invention is due to Thorne and is fully described in a paper by him.[3]

7. BEAM SWITCHING

It has been mentioned already that a minimum of three antenna beams are required to define completely the velocity vector of the vehicle. Antenna designs are such that all beams required can be produced by a single aperture and with certain types of antenna these beams can be used simultaneously. There is, thus, a choice to be made between using a multi-channel system with all beams operating all the time, or a single channel system using beams sequentially. With a normal navigational Doppler system in which speeds only from say 70 to 1000 knots need be considered the choice is entirely one of economics and relative reliability. The economic factor is resolved by comparing the cost of an antenna switch and control equipment with the

additional cost of providing additional receiving and tracking channels. It is also necessary to take account of the fact that mechanical waveguide switches tend to be in the category of unreliable components although this is not necessarily true either of ferrite switches or switches employing microwave diodes; the latter, however, do tend to be inefficient. In the case of equipment required to operate to near zero and perhaps negative velocities a further complication arises in that with switching of beams, steps have to be taken to prevent the switching frequency, or harmonics thereof from falling into the Doppler band. For helicopter systems and similar applications therefore parallel beam operation is generally to be preferred.

8. DUPLEXING

With all types of radar system it is necessary to take steps to ensure that the transmitted signal does not enter and overload the receiver. This may be achieved with a duplexing device such as a ferrite circulator or system of TR switches coupled with pulse modulation of the transmitted waveform. Similar duplexing devices are appropriate whether the transmitting system be pulse coherent or self coherent. Ferrite type duplexing systems are also applicable with F.M./c.w. transmitters although with some systems it has proved difficult to achieve enough isolation by these means. With unmodulated c.w. systems it becomes necessary to use separate antennas for transmitting and receiving in this case an isolation between antennas of some 80 db being required to provide operation to 70,000 ft over Beaufort 1 sea. This figure is fairly readily achievable provided there is no coupling by a common radome.

Referring to Figs. 1–5 it will be seen that using the arguments just presented some effort can be made to categorise Doppler systems.

Dealing just with general purpose systems for transport type operations, these are generally coupled with some form of gyro magnetic compass and a simple navigation computer. Computers of both analogue and digital types have been used, the latter being usually special purpose pulse counting type computers rather than general purpose digital computers.

Computer performance lies in the range such that from perfect inputs position will be indicated to within 0.5% of distance travelled. Heading errors can be anywhere from about 0.5° to 2° depending on the system and in particular how well and frequently compass swings are carried out.

The normal performance specification for a Doppler sensor to mate with such equipment would be up to 0.5% maximum groundspeed error and up to 0.5° drift error. These are broadly the characteristics of equipment to meet the requirements of ARINC characteristic 540.

The overall accuracy of such a navigation system would be very largely dependent on the particular heading reference used but typically an error of between 1% and 3% of distance travelled should be expected with the error across track exceeding that along track.

Three typical configurations which have been used for this type of service are shown on Fig. 6 and are marked *A*, *B*, and *C*.

TYPICAL SYSTEMS

System A	F.M. c.w. Transmission	Superhet Receiver	Duplex Operation	Fixed Antenna	Pencil Beams	Sequential Operation
System B	F.M. c.w. Transmission	Superhet Receiver	Separate Antennas	Drift stabilised Antennas	Conical Beams	Sequential Operation
System C	Self coherent pulse transmission	Superhet Receiver	Duplex Operation	Drift stabilised Antennas	Conical Beams	Sequential Operation
System D	Unmodulated c.w. transmission	Homodyne Receiver	Separate Antennas	Fixed Antennas	Pencil Beams	Parallel Operation
System E	Pulsed coherent transmission	Superhet Receiver	Duplex Operation	Fixed Antennas	Pencil Beams	Parallel Operation

FIG. 6

Comparing systems *A* and *B* it will be noted that there are variations under three headings in that system *A* has a fixed duplex antenna system with pencil beams whilst system *B* has a drift stabilised double antenna with conical beams.

It would seem likely that the designer of system *A* could achieve duplex operation because of the more stable microwave conditions achieved with a fixed antenna while the designer of system *B* could not achieve sufficient isolation with the moving antenna without the use of separate antennas.

Thus system *A* would be easier to install than system *B* but it would be expected that system *B* would have better drift measurement accuracy. Both being F.M. c.w. systems would suffer altitude dependent phenomena but would be expected to provide operation well within 0.5% maximum error averaged over several minutes. System *B* having conical beams would suffer less from altitude dependent errors than system *A*.

Looking now at system *C*, this differs from system *A* in type of modulation and the ability to operate duplex. This system being self coherent might be expected to be rather more accurate than system *B* but would suffer in tolerance to roll and would not operate to such a low altitude.

In the case of higher accuracy systems intended for use with inertial quality heading references, the approach most likely to yield the required accuracy would be system *C* using the largest antenna that the aircraft can accommodate. In this case a very narrow spectrum would be obtained without extensive frequency stabilising systems while drift stabilisation of the antenna would provide high accuracy of drift measurement and would make drift accuracy independent of sea state. Calibration change of ground-speed with sea state would also be minimised by the narrow antenna beams. The origin of sea state errors is discussed by Berger[2] and by Gray and Moran[4] and so will not be treated here.

There is one class of Doppler system not yet considered and that is the system for helicopters. Such a system has the requirement of measuring velocities down to zero and through zero to negative values. This adds considerable complexity and demands the use of coherent systems so that the vertical component of the velocity may be obtained.

Systems *D* and *E* on Fig. 6 are typical solutions. In a comparison of these systems, *D* using a homodyne receiver would require more transmitter power and being unmodulated would require separate transmitting and receiving antennas with separate radomes.

System *E* would be much more complex in respect of transmitter design but could use a single antenna.

9. FORTHCOMING DEVELOPMENTS

There are one or two techniques now becoming available which are likely to influence profoundly future generations of Doppler system. Of these perhaps the most significant is the ability which now exists to generate a significant amount of power at microwave frequencies by frequency

multiplication from a crystal controlled driving oscillator. The device which makes this possible is the varactor diode.

Typically it is now possible to generate about 100 mW of power at a frequency of 12,325 Mc/s by multiplying up from a crystal controlled oscillator operating at 92.5 Mc/s. The character of the resulting output is one of very high stability and the frequency is very stable against change of load impedance. The sideband noise near the carrier which is important for low speed Doppler application appears to be rather worse than the best klystrons at the present state of the art but is very dependent on design of the multiplier unit.

It is, furthermore, possible to modulate these devices both by frequency modulation and to provide coherent pulse output. The latter can be achieved by a combination of gating at the low frequency oscillator output and adding a diode switch at the microwave output. This is a much simpler method of achieving coherent pulse transmissions than the method described earlier using klystron amplifiers.

It seems likely that with the advent of these microwave sources there will be produced systems of unmodulated c.w. type for general purpose use and for the most stringent requirements pulse coherent systems with azimuth stabilised antennas.

Such systems will probably incorporate a considerable degree of micro-miniaturisation resulting in very lightweight and it is hoped reliable equipment. Some work is also justified in attempting to minimise antenna weight possibly by stripline techniques.

Some work is known to have been carried out to reduce to negligible proportions the change of calibration with sea state. One approach is akin to monopulse radar techniques in which two beams are used in each direction which are slightly displaced from each other in the depression plane. Thus two overlapping spectra are produced and the cross-over point of the two is tracked. This is much more stable than tracking a spectrum which suffers distortion due to change of scattering coefficient. However, it is the Author's opinion that the considerable extra complexity introduced is not worthwhile in view of the possibly quite large errors due to sea surface movement which remain uncorrected. To take an example, a conventional system operating with a maximum error of $\pm 0.2\%$ over land might have a 1% underread over average sea. A simple land/sea switch can be fitted to change the calibration over water by 1% and the spread over water is then likely to be about $\pm 0.4\%$ with change of sea state. However error due to water surface movement has been shown to be about 1/5 of the surface wind velocity so that with a surface wind of 25 knots which is by no means uncommon, an error of 4 knots will occur which at 300 knots cruising speed will represent an error of 0.75%. Thus with the simple system incorporating a land/sea switch, the maximum error over water would be 1.15% while with the sophisticated and complex double beam system the error would still be 0.95%. This implies a lot of complexity for a marginal improvement.

It may be of interest to conclude this paper by giving some indication of the progress made in Doppler development over the last 10 years. Equipment

in service about 10 years ago could provide groundspeed with an error of the order of 0.1% and drift angle with an error of the order of 0.1°. The weight of such equipment was about 220 lbs.

Typical Doppler systems today weigh between 60 lbs and 100 lbs and have errors ranging from 0.1% groundspeed and 0.1° drift angle to 0.5% groundspeed and 0.5° drift.

Systems now under development are expected to weigh about 30 lbs with systems for limited applications weighing even less.

There has thus been significant progress in weight reduction over the past 10 years but no significant movement in accuracy figures. It seems probable that the figures quoted for basic accuracy will not be improved upon significantly but there are good prospects for better overall system accuracy as better heading references become available and as digital methods of computation gain popularity.

10. ACKNOWLEDGEMENTS

The Author wishes to thank the Directors of Decca Radar Limited and the Decca Navigator Company Limited for permission to publish this paper.

DISCUSSION

G. E. Beck: Would Mr. Gray care to forecast what rate and what reliability a Doppler equipment would have in five years time, that is a Doppler equipment using the varactor techniques you mention and also using microminiature techniques to the fullest possible extent?

T. Gray: You are asking me to stick my neck out. I think that by looking at what we are achieving to-day, and at the best being achieved in other fields we can make a useful estimate. I don't know whether Mr. Beck would contradict me here or not but I would think that the current performance of Doppler systems for reliability in a reasonably benign aircraft environment is probably from two to four faults per thousand hours. I would think that by introducing all the best measures that are available in a few years time one ought to get into the bracket of 1 fault in 2,000 hours, I would not care to go beyond that. But this sort of mean time between failures is already being achieved with certain classes of equipment, and there would seem to be no reason why that sort of figure should not be achieved with Doppler radar. I think that by the next year or two we will achieve a system to do the full job that we have been describing in perhaps 30 lbs or so, without taking any special measures to lighten antennae, that is using the hollow tube type of hardware. I have a slightly open mind as to whether there is justification for diverting effort into antenna lightning techniques but this is the only area where one I think can make a significant contribution. It is already perfectly possible to design out the pieces which in the last generation of equipment contributed so much to the weight, such things as analogue computing

elements with synchro-servo motors and so on. This accounts for the drop in weight between perhaps 60 lb of present equipment and the 30 lbs of the one which I now postulate.

REFERENCES

1. Freid, W. R., Principles and performance analysis of Doppler navigation systems, *Trans. I.R.E.*, **ANE-4,** 4, 176–96, 1957.
2. Berger, F. B., The design of airborne Doppler velocity measuring systems, *Trans. I.R.E.*, **ANE-4,** 4, 1957.
3. Clegg, J. E. and Thorne, T. G., *Doppler Navigation*, published by the Institution of Electrical Engineers, March 1958.
4. Gray, T. and Moran, M. J., Decca Doppler and air navigation, *Brit. Commun. and Electronics*, **5,** 10, 764–771, 1958.

CHAPTER 14

SCANNING BEAM GUIDANCE FOR APPROACH AND LANDING

J. E. WOODWARD
Airborne Instruments Laboratory,
Deer Park, Long Island, New York, U.S.A.

An account is given of a radio guidance system for aircraft which is a combination of the usual ILS and a precision approach radar. Two ground stations are used in order to give control in azimuth and glide angle. These are sited one at each end of the runway. A similar system could be applied to landings on aircraft carriers.

1. INTRODUCTION

Radio guidance of landing aircraft has been under development for more than three decades but has yet to reach operational fulfilment. During World War II, radio guidance techniques suitable for low approach were brought into practical use in the now familiar ILS and GCA systems. These two systems have been improved over the years without changing in basic design and capability.

The difficulty in closing the gap between approach guidance and complete landing guidance is evident from the development effort in this area over the past 10 to 15 years. A variety of techniques has been studied or tested during this time, with many of them having been developed to the point of demonstrating the technical capability of landing aircraft. One group of techniques, based on the use of radio altimeters, has blended into an approach that is now being given serious consideration for use in a landing guidance system.

The technique now under consideration for evaluation as an advanced landing guidance system is actually a combination of some of the best technical and operational features of ILS and GCA. This new system, currently being developed for the FAA, preserves the narrow scanning microwave beam of the GCA with its ease of siting and wide angular coverage, and provides the ILS features of air-derived data and virtually unlimited traffic capacity.

In consideration of the more demanding requirements for elevation guidance during flareout and touchdown, considerable effort has been made to develop a suitable technique for this region of flight. The ground-referenced technique that has proven most satisfactory uses narrow scanning microwave beams. Scanning-beam guidance systems, such as FLARESCAN, have demonstrated the ability to supply the accurate elevation data required for flareout and touchdown. The FLARESCAN technique has been used with excellent results in test programs in the United States and France.

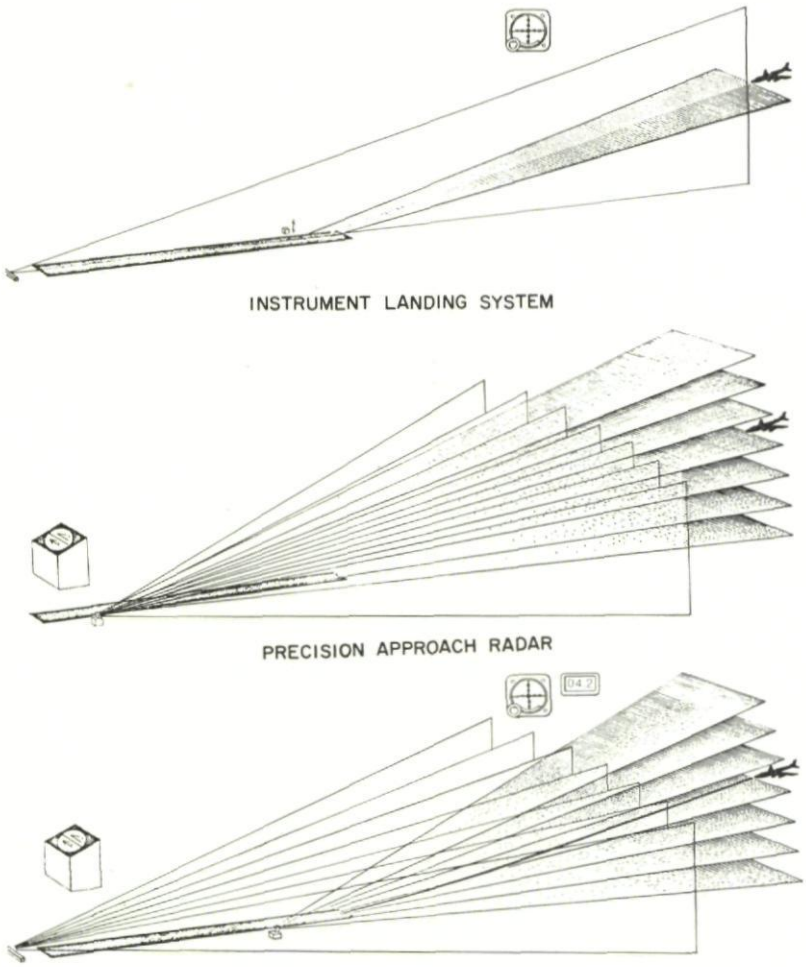


FIG. 1. *Combination of ILS and GCA functions.*

The ability to provide flareout and touchdown guidance as well as other advantages led to the selection of scanning beam techniques as the basis for a new guidance system for approach and landing. This system, known as the Advanced Integrated Landing System (AILS), provides full air-derived landing guidance (including precision DME) as well as ground-derived radar (GCA) surveillance information. In normal civil applications, the ground-derived radar information is used to monitor approaching aircraft; GCA low-approach capability provides a backup. The GCA function is not sacrificed however. In fact, the latest techniques are used to provide full capability for low-approach guidance. Figure 1 shows how ILS and GCA are combined in the new system.

This paper discusses the overall design and important technical details of AILS. It also covers several other applications of scanning-beam techniques that are under consideration.

2. ADVANCED INTEGRATED LANDING SYSTEM (AILS)

Figure 2 shows the basic elements of AILS. The system consists of an azimuth and elevation ground station, a ground display site, and an airborne

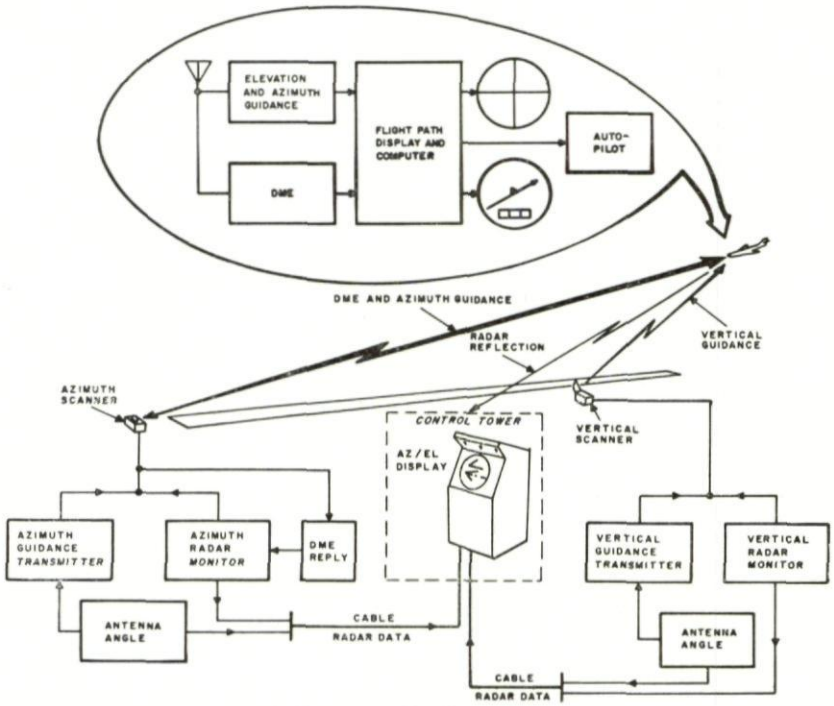


FIG. 2. *Elements of landing guidance system.*

unit. The two ground stations are situated in the same general manner as the ILS, with the azimuth station located on the extended runway centerline at the stop end of the runway, and the elevation station located alongside the runway in the touchdown region, but some 1,000 to 1,500 feet behind the intended touchdown point rather than in front. The DME transponder is located at the azimuth site. The DME system delay is adjusted to provide zero range at a point on the runway opposite the vertical scanner. This arrangement provides data in the aircraft and on the ground in the most directly usable form.

The method used to provide air-derived data is based on techniques of the Flarescan system. The radar uses the latest techniques, including circular

polarization, MTI, and bright display. The DME is tailored to this application and takes advantage of existing angle data decoder circuits in the airborne receiver for range decoding.

One of the basic system requirements as specified by the FAA was to maintain a common reference for both air- and ground-derived data. This requirement is met by using the same scanning antenna for both beam-guidance (air-derived) data transmission and the radar function. The microwave beam transmitted by this rugged antenna has a constant boresight angle with respect to the antenna. A precision angle data pickoff ensures accurate determination of the antenna pointing angle. This is important because the guidance information in the cockpit and on the ground have a common reference.

The requirement for using the same antenna for airderived guidance and for radar imposed some definite restraints on the overall system design. The most important was the choice of mixing air-derived guidance and radar on a pulse-interlace or scan-interlace basis.

Using the same transmitted pulse for both air-derived guidance and radar (pulse interlace) provides both functions on each direction of antenna scan. This minimizes the scan rate needed for a given data rate. However, it also presents a serious problem because the requirements for air-derived guidance are quite different from those for radar. At the beamwidths and data rates needed for this system, the PRF required for good air-derived guidance would result in a radar range between pulses of about 4 miles. This is far from adequate for the radar function.

The use of scan interlace, where radar and air-derived guidance are transmitted on succeeding scans, enables the parameters of each function to be individually optimized. A possible disadvantage is that twice the antenna scan rate is needed for the same data rate obtained by pulse interlace. However, a simple and rugged scanner has been developed that will provide a scan rate of 5 c/s. This gives the required 5 samples per second data rate for air-derived guidance and a more than adequate data rate for radar. With the angle data pickoff, the angular registration between the radar and guidance scans will be better than required.

An extension of the scan-interlace technique is used to put all ground-to-air functions for a given runway on a single RF channel. The scan cycles from the azimuth and elevation scanning antennas and a DME time cycle are time-shared so that only one function is transmitted at a time. Figure 3 shows the method of time sharing. The time required for a complete antenna scan cycle is divided into six periods. The antennas are made to scan mechanically well beyond the angles of active coverage (± 11 degrees versus ± 5 degrees of active coverage). The time it takes one antenna to overscan the active sector provides sufficient time for two other transmission periods. This scheme requires that the scanning antennas operate not only at the same frequency but at a fixed phase relationship.

Two of the six time periods are used for guidance (azimuth and elevation), two for radar (azimuth and elevation), and one for DME. The sixth period is reserved for possible future addition of a second elevation site or to provide

additional DME time. Locating an additional elevation guidance site at the glide-slope aiming point would simplify glideslope guidance. More important, it would permit elevation position to be determined in the aircraft from three combinations of data (DME and two angles). Thus, an error in any one data source could be detected during approach and landing and could be completed using the other two sources. This form of system redundancy is believed more desirable than the usual equipment redundancy.

The choice of an operating frequency for the guidance system was based on a number of opposing factors. The principal factors were: the deleterious effect of rainfall as the frequency is increased, and the unmanageable size of

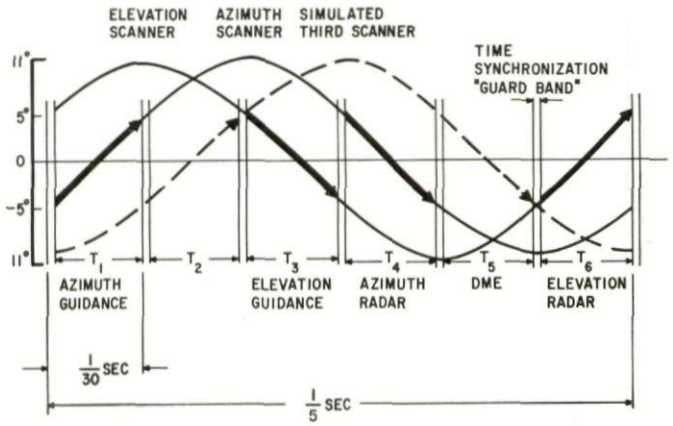


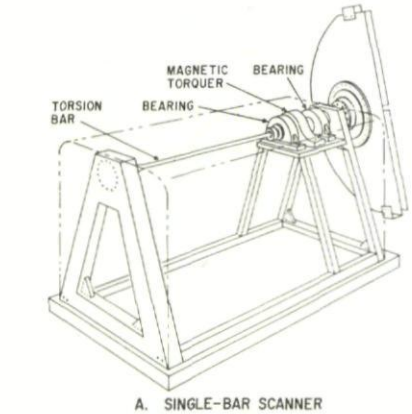
FIG. 3. *Scanning and guidance cycle.*

the scanning antenna aperture as the frequency is reduced. The decision was made to operate the system at K_u-band, between 15.4 and 15.7 Gc/s. At this frequency, guidance requirements can be met at ranges well in excess of 15 miles with a moderate transmitter peak power of 2 kW. A transmitter peak power of 130 kW is used for radar tracking. Scan interlace makes possible the use of two different transmitters.

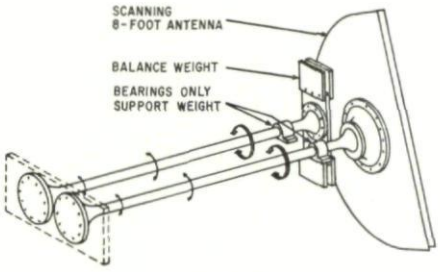
An analysis of radar range capabilities has shown that useful radar tracking to a range of about 10 miles can be expected with a heavy rainfall of 10 mm per hour over the entire path. Although a decrease in operating frequency (to X-band, for example) can slightly increase radar range performance, it is not a major factor in system design. The design recognizes radar tracking as a monitor and backup function which by itself cannot provide satisfactory guidance to touchdown. If extended tracking range is desired, it can be best provided by secondary radar. The airborne DME transmitter offers the future possibility of secondary radar plus the benefits of a great increase in range at moderate power levels and the elimination of clutter and angle jitter or scintillation on the azel display.

The scanning antenna is a crucial element in the system design because it is common to both radar and beam guidance. A simple and reliable

mechanical scanning technique has been developed and thoroughly tested. Two versions of this scanner are shown in Fig. 4. The antenna (a thin pillbox) is mounted on the end of a steel bar with the other end of the bar held stationary forming a torsional pendulum. If the antenna is rotated on the axis of the bar (twisting the bar) and quickly released, it will oscillate at a frequency determined by the spring constant of the steel bar and the



A. SINGLE-BAR SCANNER



B. DOUBLE-BAR SCANNER

FIG. 4. *Elevation scanners.*

moment of inertia of the antenna. For fixed-base operation of AILS, the stationary end of the bar is anchored to the earth (Fig. 4a). For a transportable or mobile system, an additional torsion bar and dummy antenna are used (Fig. 4b). This system, as used in the present Flarescan ground station, balances the twisting force of one bar by matching it with the out-of-phase force of the other bar, so that no force is transmitted to the ground. Because the losses from windage and bearings are small compared with the stored energy in the oscillating system, the Q of the oscillating system is very high. The necessary driving energy is supplied by a magnetic torque motor pulsed in step with antenna motion to form a self-resonant drive system. By exciting the system at its natural resonant frequency, the average torque or energy required to keep the resonant system running is held low. The

torque motor is mounted coaxially around the torsion bar to eliminate gears and linkages.

In the full landing system, two such scanners must be operated at the same frequency and at a fixed-phase relationship. This is done by operating each scanner at its natural resonant frequency and continuously tuning the resonant frequency of each scanner to a precision crystal-control frequency reference. This tuning is done by automatically adjusting the radial position of a pair of weights on the antenna to control its moment of inertia. Without tuning, changes in resonant frequency as a function of temperature (the only significant disturbing effect) are slight and occur at a very slow rate. Thus, compensating for these changes is not difficult.

With the two scanners accurately tuned to the same scanning frequency, the elevation scanner must be phase locked to the azimuth scanner. This is done by sensing any phase error between scanners and shifting the phase of the magnetic driving pulses relative to the scan cycle of the elevation scanner. This phase shift produces a slight detuning or frequency shift of the scanner and reduces the phase error.

The control circuits for self-resonant drive, tuning, phase lock, and amplitude control are all digital and operate from a precision 10 kc/s timing reference. The digital timing circuits, in conjunction with angle data pickoffs on the antenna, also supply timing gates to the rest of the system for interlacing the various functions during a scan period. The scanner drive and control circuits have been thoroughly tested and operate as desired.

3. AIR-DERIVED GUIDANCE

The AILS system will provide all of the information required to make a complete instrument approach and landing. This information consists of precision elevation angle, azimuth angle, and distance from a point on the runway. The air-derived data may be processed in a number of ways for display in the cockpit for manual flight control and/or application to an autopilot for automatic approach and landing.

Because of the various functions of the landing guidance system operated in time sequence into a single airborne receiver, the receiver must be able to identify each function as it is received. The signals being transmitted from ground to air are all on a single RF channel in the 15.4 to 15.7 Gc/s band. Ten different frequency channels are available to eliminate potential interference from equipments on nearby runways. The various functions from a given runway installation are identified by pulse coding. A simple two-pulse code with different pulse spacings is all that is needed for the identity function. At the same time, the code provides an appreciable measure of immunity from interference. The following pulse-coding assignments have been selected:

DME reply	8 μs
Azimuth angle	10 μs
Elevation angle	12 μs

The DME air-to-ground interrogation will be done on a single RF channel. The identity of the runway equipment being interrogated will be determined by pulse coding using 10 two-pulse codes, ranging in pulse spacing from 15 to 33 μs. Each code will be associated with an RF channel so that, when the airborne receiver is tuned to a particular channel, the associated DME interrogation code will be transmitted. The single-pulse transmission of the ground radar function, although on the same RF channel as the beam guidance information, will be rejected in the receiver.

3.1. *Elevation and Azimuth Angle Transmission*

The technique for angle data transmission developed for FLARESCAN will be used in AILS with no basic changes. The minor changes and adaptations required for use in both elevation and azimuth guidance will not affect the proven performance capability. The FLARESCAN technique had,

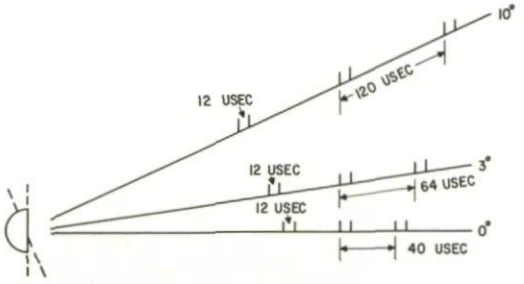


FIG. 5. *Elevation angle data encoding.*

in fact, been developed to accommodate an expansion to a two- or three-beam guidance system. Therefore, the scan rate and timing can easily accommodate the additional functions.

The angle data transmission from the ground consists of a series of pulses (pulse-pair groups in this case) in which the spacing between the pulse groups represents the scan angle of the antenna at that instant. As shown in Fig. 5, 10° of antenna scan is represented by a change in pulse group spacing of from 40 to 120 μs. Thus, the pulse group spacing has a variation of 80 μs to represent 10° of antenna motion, or 8 μs per degree of motion.

As the antenna scans, a photocell views an engraved scale attached to the antenna. The output of the photocell is a pulse or mark for every 0.01 degree of antenna scan. A total of 1000 angle increment pulses are counted and stored in the angle encoder. The encoder translates the stored scan angle information into the variable-spacing angle data code. Using a crystal oscillator as a precision timing reference, the output of the angle data encoder represents the instantaneous angle of the antenna to an accuracy of $\pm$0.01 degree.

As the encoded beam scans over an airborne receiver, the receiver must first decode identity to determine which function is being received. This is first done on the sidelobe level or low skirts of the beam as it approaches the aircraft. This sets up the receiver circuits for the proper function and

establishes the proper AGC level before angle data is taken from the beam. The angle decoder in the receiver decodes the angle message represented by each succeeding pulse group and determines the angle representing the peak of the beam by taking the center of gravity of the angle messages received. The angle data received on each beam passage is used to update the appropriate memory so that continuously updated information is available for each function. With a scale factor of 8 μs per degree, it is not difficult to

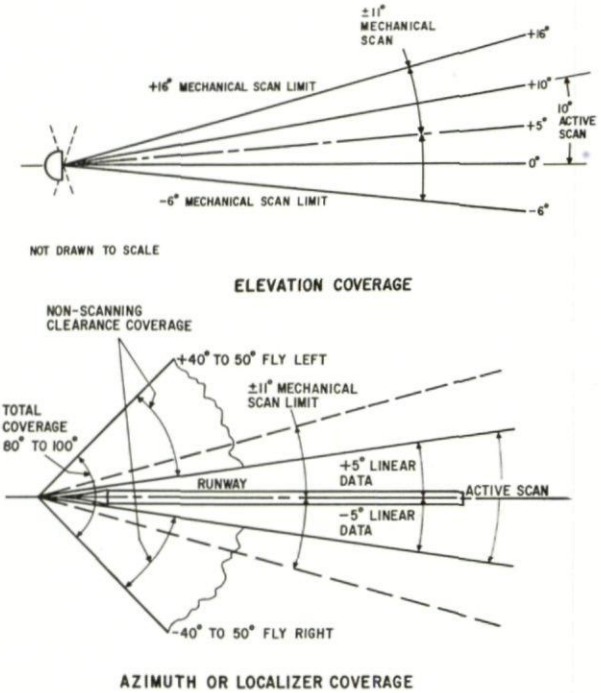


FIG. 6. *Elevation and azimuth guidance coverage.*

obtain the required short-term stability in the analog decoder. An internal automatic calibration routine is performed once each scan cycle to ensure long-term stability.

The same angle encoding and decoding technique is used in the azimuth or localizer scanner, except that the midpoint of the scan is used as the zero reference. The angle memory provides a measure of deviation from the centerline of the runway, providing up to 5° of fly-right or fly-left guidance.

Figure 6 shows the coverage of the beam guidance information in both elevation and azimuth. The elevation coverage is 0 to 10°. Actually, the encoding will be available for about 0.25° below zero, and the lowest angle at which data are transmitted is adjusted at each site to provide optimum low-angle coverage. Also, in practice, the elevation scanner is adjusted so that an encoded angle of zero degrees corresponds to a line parallel to the

runway at the point of intended touchdown. This may differ slightly from horizontal, depending on the grade of the runway.

In the azimuth scanner, angle data are again transmitted over the total angle of 10°, or ±5° from the centerline of the runway. Although a linear or proportional data coverage of ±5° is adequate, additional angular coverage is needed to inform the pilot that he is off course to the right or left of the centerline. This clearance coverage is specified in the latest ICAO specification to extend to 35° on either side of the centerline. Clearance coverage is provided by two nonscanning antennas (Fig. 6). A short transmission of about 3 ms (equivalent to the dwell time of a scanning beam on an aircraft) is transmitted from each of these nonscanning antennas in sequence. One transmission is sent just before and the other just after the active scan of azimuth guidance.

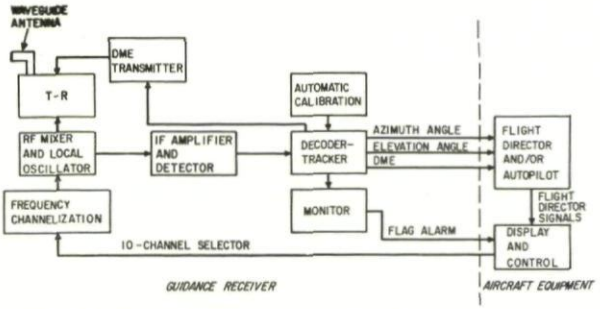


FIG. 7. *Simplified block diagram of airborne equipment.*

An aircraft flying in the clearance coverage of the system will interpret these short transmissions in the same way as a scanning beam. The data transmitted in each clearance beam will be interpreted by a receiver located within that clearance area as an angle greater than 5° on the appropriate side of the centerline. The signal level required in the coverage area must be great enough to overpower sidelobe energy from the scanning antenna, but not strong enough to overpower the main beam of the scanning antenna. The clearance antenna has about 1% of the gain of the scanning antenna. Therefore, the required signal level for clearance transmission is obtained by using the high-power radar transmitter. The ratio of antenna gains between beam guidance and clearance transmission is offset by the ratio of transmitter powers between the beam guidance and radar transmitters.

Figure 7 is a simplified block diagram of the airborne equipment used to receive beam guidance. A superheterodyne receiver provides the necessary selectivity. A crystal controlled, solid-state local oscillator ensures receiver frequency stability. The frequency of the local oscillator can be switched to any one of 10 available channels.

The received angle data are processed in the decoder-tracker and the resultant angles are stored in an appropriate output memory for use by the flight director or autopilot. The decoder-tracker sorts out or decodes the function being received, decodes the angle-data code, and determines the

angle representing the peak of the beam as it passes over the aircraft. The decoder-tracker circuits are used in turn for the two angles and for the DME function. The previously mentioned automatic calibration routine is performed at an angle of 0° to ensure accuracy at low elevation angles and at an angle of 5° representing localizer on-course. This automatic calibration against a crystal-controlled timing reference makes possible the use of a simple analog decoder to obtain the precision decoding required.

3.2. *Distance Measurement*

Associated with the air-derived angular guidance is a precision distance measuring system. The distance of the aircraft from a selected point on the runway is determined by measuring the time required for a pulse to travel from an airborne interrogator to a ground transponder and back to the aircraft. Fortunately, the most difficult function of the DME system (the measurement of time delay) can be readily accomplished by the angle data decoder. It was decided to place the DME transponder at the azimuth scanner site because the distance to a point on the runway could be measured directly without having to compensate for the slant-range error that would result from an offset site.

The system design philosophy of separating each function into an appropriate time slot was applied to the DME function for two reasons. First, the pulse-timing requirements for DME are not compatible with either the guidance or radar functions. Second, it is not desirable to operate the DME transponder on the narrow scanning beam antenna. Use of the narrow beam would compress most of the DME activity into the time of one or two beamwidths of azimuth scan. Most of the aircraft within the coverage of the system would be lined up on approach or on the runway, or would be grouped in a departure queue near the runway threshold. Although it may be possible to handle this concentration of aircraft in the few milliseconds of beam passage, this would place some unnecessary requirements on equipment design, peak duty cycle, etc. By using an entire time slot of $\frac{1}{30}$ second, reasonable airborne interrogation rates can be used, and the possible loss of service as a result of transponder dead time can be minimized. An analysis of the service that will be provided for up to 10 aircraft shows that the system will be more than adequate.

The airborne interrogator will use a magnetron operating at a peak power level of 2 kW, and at an average prf of 600 pulses per second during the DME period. The basic airborne timing standard for measuring the time between interrogation and reception of a reply is a combination of the crystal-controlled clock pulse generator of the autocalibration circuit, and the same time decoder that is used for angle data decoding. The timing pulses from the autocalibration oscillator occur at exactly 40-μs intervals. The interrogations, with the appropriate pulse-pair spacing, are transmitted in step with one of these pulses, and the second succeeding pulse is considered zero range, providing a system dead time of 80 μs.

The decoder covers a basic spread of 80 μs, which corresponds to a range of about 7 miles. At longer ranges, the start of the decoder is delayed an

additional 80 μs, which corresponds to the fourth clock pulse following interrogation. The decoder will then operate from about 7 to 14 miles. In fact, the range is extended by operating the decoder beyond its normal angle decoding range in the second sector to provide a total DME range of about 17 miles. The required DME accuracy of ± 100 feet or 1% of range is compatible with the required angular accuracy in the operation of the airborne decoder. The airborne interrogation rate is controlled by an unstabilized oscillator operating at about 600 c/s to minimize interference between aircraft.

The transmitting and receiving antennas for the ground transponder provide coverage that matches the angular volume of the system. As for azimuth clearance signals, the high-power transmitter is used to offset the gain of the reply transmitting antenna. The transponder is located at the stop end of the runway and the requirement that zero range be referenced to the elevation site means that the transponder dead time must be tailored to each installation. The transponder dead time is essentially 80 μs less the propagation time from the zero reference to the transponder and back. This dead time is controlled by an autocalibration technique similar to that used in the airborne receiver. A test interrogation is sent through the transponder and the delay time of the reply is checked against a crystal oscillator timing standard. The delay will be continuously corrected to minimize error. The signal level of the test signal will represent an interrogator near runway threshold to optimize calibration at this range.

The autocalibration cycle will take place at the beginning of the DME time period. This serves the additional function of alerting the airborne receiver to the beginning of the DME period. The prf of the DME autocalibration function, which will have a duration of about 3 ms, will be about 5000 pulses per second. Reception of this unsolicited burst in the airborne receiver will initiate a $\frac{1}{30}$ second DME interrogation period.

4. GROUND-DERIVED GUIDANCE

Ground derived guidance is provided from both the elevation and azimuth ground sites by radar skin tracking. This GCA radar function is time-shared on the scanning antenna with the air-derived guidance function a each site. This provides a radar data rate from each site of five samples per second. Other features of the radar are circular polarization to minimize the effect of rain clutter and MTI to minimize the effect of both rain and ground clutter. The precision azel display will use a direct-view bright-display cathode-ray tube.

Siting the azimuth radar on the extended runway centerline offers several advantages. Because the aircraft will approach on a radial line, no offset cursor with its alignment problems will be required. Another advantage is the ability to display a radar picture of the landing runway and its approaches. A possible display of this information is shown in Fig. 8. The ability to survey the runway is considered of fundamental importance for all-weather landing capability. A means of positive checking for runway

occupancy is necessary both from the standpoint of safety and high traffic capability. This function is provided by Airport Surface Detection Equipment (ASDE) at some airports. However, the runway surveillance feature of the landing system is considered important as a substitute when ASDE is not available and as an adjunct to ASDE when it is available.

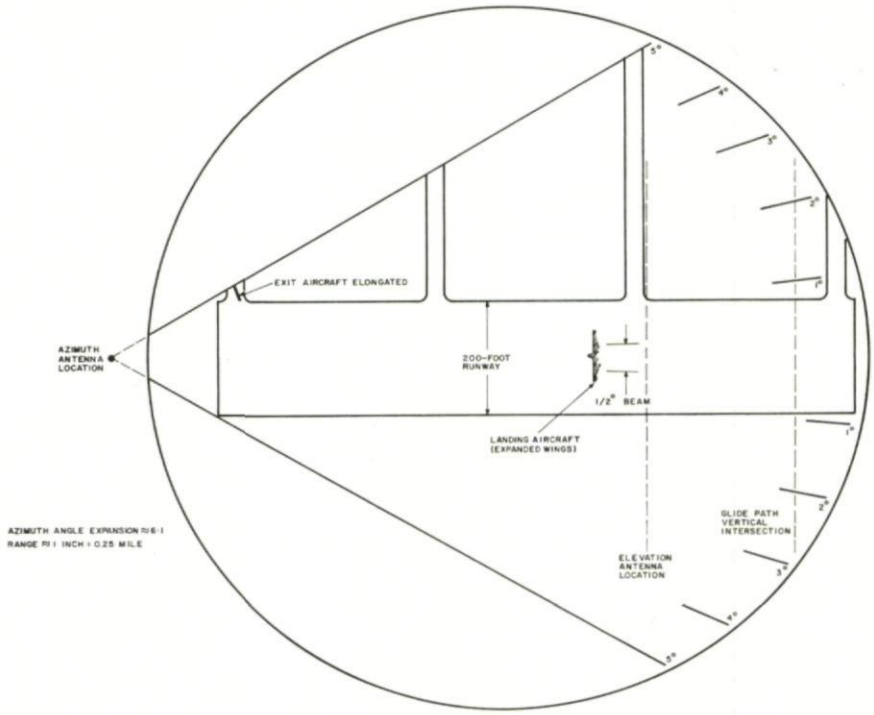


FIG. 8. *Runway display with 2-mile range.*

The radar transmitter has a peak power output of 135 kW, a pulse length of 0.2 μs, and an average prf of 5500 pulses per second. This transmitter is also used to transmit azimuth clearance guidance and DME transponder replies.

Both the elevation and azimuth scanning antennas will be circularly polarized. The elevation antenna will have a 0.5° half-power beamwidth in elevation and an azimuth-plane beamwidth of about 20°. The azimuth site antenna will also have a 0.5° beamwidth in the direction of scan, but will have a shaped beam in the vertical plane. This is necessary for both radar and beam guidance to minimize the ground lobing and attendant nulls in the vertical coverage that would be present if significant energy were allowed to strike the ground. The primary elevation pattern of the elevation antenna will be about 3° between half-power points, providing a sharp cutoff on the lower side of the pattern. It will be tapered on the upper side of the pattern to provide cosecant-squared coverge to at least 10°. The

shaped reflector needed to achieve this requires a heavier antenna (than that required for elevation) which is more difficult to scan. However, a torsion-bar scanner design is practical with this antenna.

The MTI incorporated into the radar design is a noncoherent system that uses double cancellation and has staggered prf to eliminate blind speeds in the range of interest. The first MTI blind speed would occur in the middle of the range of interest—about 120 mph. The use of noncoherent or clutter-coherent MTI was chosen to optimize discrimination against rain return.

The precision angle data pickoff used for the air-derived guidance function is also used to transmit angle data to the azel display. The angle increment pulses are sent to the display site where they are counted in a digital counter that is also used as a digital-to-analog converter. Thus, a DC voltage accurately proportional to scan angle is available to drive the azel sweep circuits. Also available from the counter are precision angle data pickoffs for generating angle marks on the display. This method of transmitting angle data to a remote display provides great accuracy, minimizes the possibility of interference, and is expected to be a decided improvement over present GCA techniques.

The accuracy that can be obtained from the radar display depends on a number of factors. The more important of these are the antenna beamwidth, the accuracy of angle pickoff, angle transmission and display, the scale factor of the display, the ability of an operator to determine the center of the target, and angle jitter of the apparent target. Of these, angle jitter of the target is the chief limiting factor at the critical distances of final approach and flare.

The apparent radar target shifts from one part of the aircraft to another as a result of very slight shifts in aircraft attitude. The apparent radar target may shift between engine nacelle in azimuth, and between tail cap, cockpit, nose, and wheels in elevation. There is no possible correction for this in the radar. It can be corrected only by the use of a corner reflector on the aircraft or by the use of secondary radar with a transponder in the aircraft. Secondary radar has been mentioned as a future expansion capability of this system, making use of the DME interrogator-transmitter.

The AILS that is being constructed for the FAA will be ready for delivery in February 1965. It is scheduled to be installed at the FAA National Aviation Experimental Center (NAFEC), Atlantic City, N.J., for a test and evaluation program.

5. OTHER SCANNING-BEAM GUIDANCE APPLICATIONS

The same scanning-beam techniques that have proven the best choice to meet the demanding requirements of AILS are also ideally suited to a number of different and often less demanding applications. These applications range from a transportable version of AILS for heavy military aircraft to a man-pack system for counter-insurgency warfare. The wide coverage of precision angle measurement makes this technique very versatile since it can handle any approach glide angle that is required. This feature makes

the technique easily capable of handling helicopters and V/STOL aircraft as well as conventional fixed-wing aircraft.

The equipment used to transmit angular guidance to an approaching aircraft can be materially simplified when the performance requirements can be relaxed. However, it is considered important that, by using common angle data techniques, the airborne and ground equipments of the various systems will be compatible. For instance, when an aircraft needs low-approach guidance only, it is not necessary for the airborne receiver to measure all angles, but only deviation from a preselected angle. This would simplify the requirements of the decoder and probably eliminate the need for DME.

This simplified fixed-angle receiver might then be operated in ground environments to provide low-approach guidance only. Such requirements are believed to exist for both civil and military use. Here again the ground equipment required for low-approach guidance only would be simplified. The beamwidths of the scanning beams could be increased, reducing the size and complexity of the scanning antennas. It also might be desirable to eliminate the skin-track GCA, with its high-power transmitter and complex receiver. The GCA function could still be retained by the use of secondary radar. The air-to-ground reply would use an existing airborne transmitter.

The important point is that the simplified airborne guidance receiver can work within its capability with the full-scale ground station, and the sophisticated airborne receiver can work with the simplified ground station. Thus, a common technique can be tailored to meet a variety of requirements and, because of the common features, still operate interchangeably. This flexibility will make further application of the technique desirable.

A number of applications that have been under consideration will be discussed briefly.

5.1. *Transportable Version of AILS*

The AILS system that is being produced for FAA evaluation takes advantage of the fixed-base installation. The ground equipment is housed in large cabinets and the scanners are anchored to the ground. A system to provide the same capability for military heavy aircraft must have the same performance characteristics, but can certainly be packaged more suitably for transportability. A military (or other) transportable system would be packaged self-contained in a shelter, probably complete with primary power, so that the system could be readily moved into place and put in operation. Such a self-contained, transportable system would use the dual torsion bar scanner (Fig. 4) so that no concrete base would be required. The equipment would be engineered to fit into the smallest space consistent with maintainability and operation. The digital portion of the equipment would be made of integrated circuits, providing benefits in both size and reliability.

Installation and alignment of the transportable system would not be difficult with the aid of self-contained leveling devices and telescopes. A radio link would be provided for transmitting antenna synchronizaton signals from the azimuth to the elevation site. The operating site for the GCA

function probably would be at one of the equipment sites. Here again a radio link could be used to transmit video and angle data from the other site.

A transportable version of AILS would provide full landing guidance at practically any runway that can handle heavy aircraft. Such a system could be unloaded and put in operation in a few hours.

5.2. *Lightweight System for Approach Guidance*

There are many situations in which the cost and complexity of full instrument landing guidance capability cannot be justified. This is true in both civil and military operations. However, good low-approach guidance may

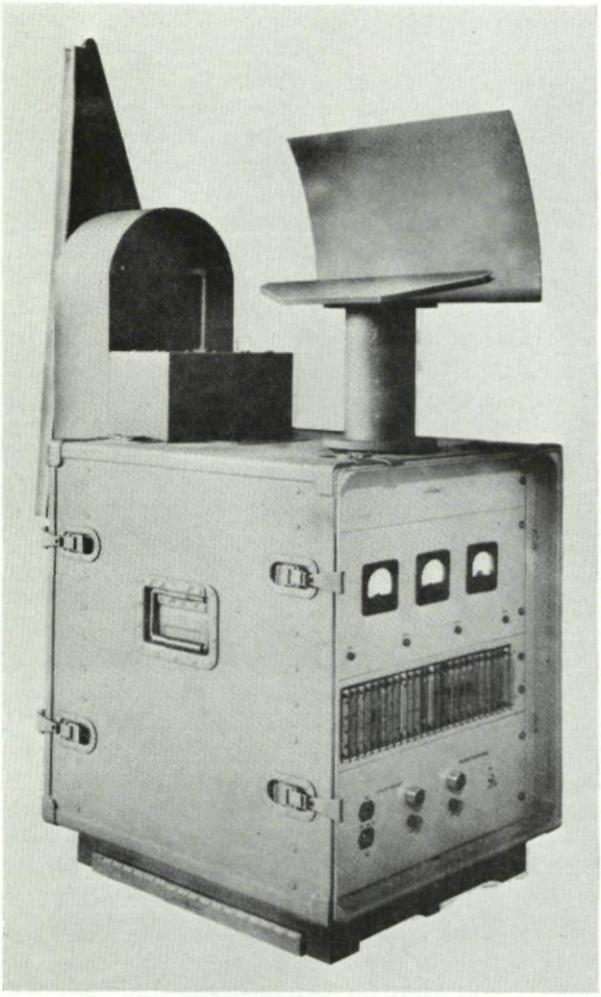


FIG. 9. *Lightweight approach guidance ground station.*

be very desirable in many such cases. This is particularly true in military operations, where it is extremely important to take advantage of bad weather and darkness for troop and supply movements. For this military use, the flexibility of cockpit-selectable approach angle is one of the strongest points in favor of the scanning-beam technique. A relatively simple ground station can provide approach guidance for a variety of aircraft types and can accommodate essentially any approach terrain in which aircraft can operate.

By relaxing the stringent requirements for flareout and touchdown guidance, the air and ground equipment can be simplified considerably. On the ground, wider antenna beamwidths permit a smaller and lighter antenna and scanner. The associated electronics are also somewhat simplified and subject to substantial reduction in size and weight. Figure 9 shows a mockup of an approach guidance ground station that is being considered. The elevation and azimuth guidance are radiated from the same point that would normally be located somewhat ahead of the approach end of the runway. The approaching aircraft would thus make a visual flare and landing, with the allowable weather minima dictated by the approach characteristics of the particular aircraft.

This system would provide air-derived or ILS-type guidance only. However, the 2 kW peak transmitter power of the ground station does permit skin tracking out to ranges of about 1 mile. This radar return can be used to automatically signal the aircraft at a preselected distance from the runway. This marker beacon feature would warn the pilot that he is at or approaching his weather minimum. At this point, he would either go visual or abort the approach.

5.3. *Approach Guidance for Aircraft Carrier*

There has been increasing recognition of the need for better approach guidance for carrier landings. Aircraft are now guided by means of TACAN and altimeters until visual contact with the aircraft carrier is reached. The pilot then uses the optical or "meatball" system for glide-slope alignment and the carrier deck for lateral alignment. In addition, he is assisted by verbal instructions from the Landing Signal Officer on board the carrier. With modern jet aircraft, the pilot has only 20 to 30 seconds to touch down when he establishes useful visual contact with the meatball. In bad weather and at night, this time is marginal for the demanding task of stabilizing on the final approach. An approach guidance system that permits stabilization on the final approach at greater range would help to decrease the percentage of waveoffs, bolters, and accidents. The equipment now in use is not adequate for the precision required. Even though a very sophisticated system aimed at automatic landing is being installed on a limited basis, an approach guidance system will still be most useful as a feeder and monitor.

A scanning-beam guidance system for this application has been devised and has received considerable interest. This system has been named C-SCAN (Carrier System for Controlled Approach of Naval Aircraft). Figure 10 shows the basic system configuration as it might be installed on a carrier.

For this application, the glide-slope angle is prescribed on board the carrier so that cockpit selection of approach angle is not desired. This permits an even greater simplification of the angle data encoding and decoding circuits As shown in Fig. 10, the proposed location for the elevation guidance transmitter provides a means of direct alignment or boresighting with the Fresnel lens optical landing system. The azimuth guidance transmitter site is shown at the stern of the carrier to provide true centerline guidance.

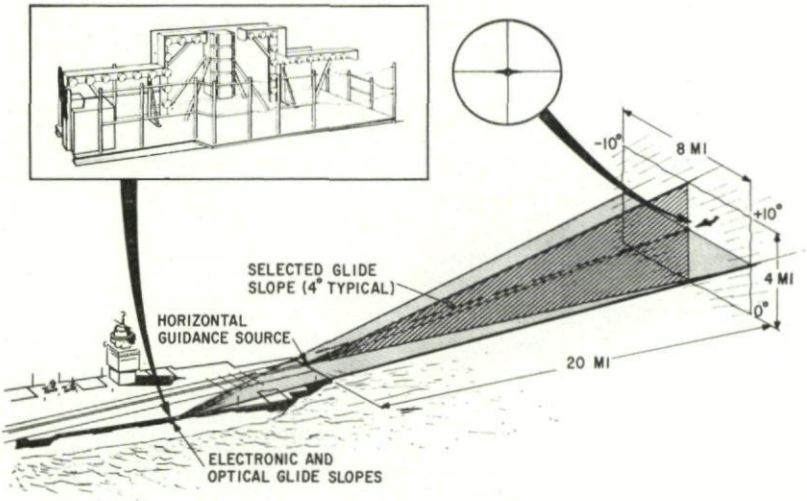


FIG. 10. *C-Scan carrier approach guidance.*

Basically, C-SCAN provides approach guidance in the aircraft cockpit out to ranges of 50 miles. This guidance would be displayed as deviation signals on a standard crosspointer meter. The displacement data would be proportional or linear over ± 1 to $2°$ from the desired path, with full-scale clearance deflection over an extended angular coverage (0 to $10°$ in elevation and $\pm 10°$ in azimuth). Glide-slope data must be stabilized against the roll and pitch of the carrier deck in the same manner as the optical or meatball system. The microwave beam from the scanning antenna is a flat sheet, matching the beam from the optical system, so that the same stabilization that is currently being applied to the optical system can be applied to C-SCAN glide-slope guidance.

CHAPTER 15

AN AUTOMATIC F.M.C.W. SURVEILLANCE RADAR

C. S. E. PHILLIPS
Royal Radar Establishment, Great Malvern, England

The paper describes a proposed fully automatic radar system which can detect and track targets both at low level and, at higher elevation angles, at ranges out to some 150 *km. A high data rate is obtained by rotating the aerial system at about one revolution per second. The elevation beam can be switched into any one of twenty three positions under the control of a digital computer. An F.M.C.W. radar is used to obtain Doppler resolutions and the elimination of fixed clutter. Progressive variations in modulation frequency enable unambiguous though inaccurate range to be obtained after two paints and for the range error to be halved after each succeeding paint. Targets can be tracked through zero radial velocity with the help of further waveform changes.*

The computer has the dual role of (1) *associating azimuth angle, elevation angle, Doppler frequency and range measurements so as to form tracks, and* (2) *controlling the aerial system, the transmitted waveform, the radar receiver and the auto-extraction equipment so as to provide the data in the most efficient manner.*

1. INTRODUCTION

In the past the human operator has played a vital role in a surveillance radar system, both in detecting and in tracking targets. Although his ability to recognise complex patterns may never be equalled by automatic machinery, his rate of operation is limited. Thus it is not surprising that, as the speed of aircraft and other targets increases, it is becoming more urgent to find satisfactory means of performing his tasks automatically. In some systems the time allowed between initial detection of a target and the reliable estimation of its track may be only ten seconds (or even less in the case of low flying targets) and this time can have a crucial effect on the overall system design. When we remember that to obtain even one wanted track, many others whether spurious or not must be considered and rejected, we are led to think in terms both of radars of higher data rates and of automatic data extraction. If a large number of tracks are to be followed simultaneously, a digital computer is required to associate the individual digitised radar outputs into tracks.

Quite apart from the advantages which accrue from automatic detection and tracking, the performance of the surveillance radar can be greatly improved if the computer is also able to control the radar itself. This control may be applied to the aerial, to the transmitted waveform and to the receiver and may be predetermined by initial programme and/or be conditional on

the state of the current data. In this event the radar and the computer must be considered as parts of a complete system.

The present paper is primarily an interim report on an experimental automatic surveillance system of this type. The ultimate objective of the work is to investigate its properties and to determine to what extent the advantages hoped for can be realised. The radar is not tied as yet to any rigid specification, but its main features are that it will operate at X-band, transmit a mean power of 2 kW via a fairly small aerial, and detect and track both long range (150–200 km) and short range (1 km) targets whose speeds can be up to Mach 3. It will be capable of dealing with targets at ground level and at elevation angles up to 60°. Range measurement will be accurate and unambiguous for all targets.

In a conventional fixed beam radar the rotation rate is usually about 5 r.p.m. This is to some extent a compromise between the desire for good detectability, which implies a slow rotation rate and the need for a high data rate to avoid confusion between neighbouring tracks. A high data rate is particularly important when one is also concerned with the low level short range target. The problem is then accentuated since the rotation rate must be increased perhaps tenfold with a corresponding deterioration in detectability. Our solution to this problem is to use an aerial whose vertical scan is electronically controlled in real time by a digital computer. A complete detection scan may take several rotations at different elevations, but not necessarily covering the full range of elevation angles. If required the detection scan time can be reduced by using elevation beams in parallel. At the same time target tracking can be performed by extrapolating data previously supplied to the computer, which directs the full power of the radar to the estimated elevation. The detection scan is thus interrupted to obtain data on established tracks and to confirm initial detections. The tracking rate depends on how often the computer signals the aerial, but is limited to the rotation rate of the aerial. The detection scan time, on the other hand, is not tied to the tracking rate and can be optimised for detectability.

A favourable feature of the proposed automatic system is that it deals with the question of clutter, since the detection and tracking of low flying targets are important requirements. Hitherto, experiments have largely consisted of the addition of automatic extraction and digital computing equipment to existing pulse radars. These radars unfortunately produce vast quantities of data from fixed objects especially when searching at low elevation angles. Since the objective of the surveillance system is to derive tracks of moving objects, these data are of no value and must be eliminated in the computer. Two problems arise in this connection, namely (a) targets within the same range bracket as a piece of clutter tend to be eliminated, i.e. they are not "seen," and (b) the computer is in danger of being overloaded. It is true that the first of these problems can be reduced by employing a wider transmission bandwidth (i.e. by shortening the pulse length either directly or by pulse compression), but this only serves to increase the data handling problem.

For these reasons we have discarded almost completely the attempt to separate targets on a range or time delay basis in favour of a narrow band range measuring F.M.C.W. radar which initially resolves targets in terms of their radial velocities. All fixed clutter returns, whatever their range, are segregated into one channel and can thus be eliminated. As the track is built up in the computer the waveform will be varied so as to improve the range measurement and to increase range resolution. (The increased range resolution is primarily required to improve the crossing-track performance of the radar.) An alternative solution might have been to use some form of variable p.r.f. pulse Doppler radar, but the F.M.C.W. radar was preferred on the grounds of relative simplicity of receiver and ease of controlling the waveform. All fixed waveform solutions, on the other hand, have serious disadvantages arising from the well known range-Doppler ambiguity problem.

2. SYSTEM DESIGN

The objective is to design an experimental system in which the radar, the automatic extraction equipment and the computer form parts of a unified automatic surveillance system. In this system the computer has two main roles to perform: (1) to operate on extracted radar data so as to give out the required information e.g. tracks, future positions, possible collision points, data to operate G.W. systems etc.; and (2) to control the radar and auto-extraction equipment so as to provide the radar data in the most efficient manner, e.g. directing the transmitted power in elevation, sequencing the ranging etc.

2.1. *The Transmitter and Modulator*

The main elements of the proposed system are shown in Fig. 1 and of the radar in Fig. 3. On the left of both figures is shown a C.W. transmitter which is capable of producing a very pure carrier wave of the required power and frequency. The C.W. output passes via rotating joints and a modulator into an aerial system rotating at a fixed maximum speed of one revolution per second. The C.W. transmitter is frequency modulated sinusoidally either directly ("internal modulation") or externally via the modulator shown. The "internal modulation" mode of operation provides strong rejection of nearby clutter and the radar is of the F.M.C.W. type.[1–4] When the external modulator is used, the F.M. is used merely as a convenient way of obtaining two frequencies and the radar is of the Two-Frequency type.[5] The deviation applied is about 50 kc/s in the first case and about twice this amount in the second case.

The modulation frequency applied is controlled by the computer which can switch on a pair of crystal oscillators whose frequencies ω_1 and ω_2 are such that $\omega_1 + \omega_2$ is constant and $\omega_1 - \omega_2$ can take up one of the following seven frequencies: 64, 80, 84, 85, 86, 88, and 96 kc/s. The frequency $\omega_1 - \omega_2$, shown for example as 96 kc/s in Fig. 3, is divided down by either 2, 4, 6 or 8, as determined by the computer, so that the final modulation frequency lies in one of the bands 32–48 kc/s, 16–24 kc/s, $10\frac{2}{3}$–16 kc/s, and 8–12 kc/s. In

Fig. 3 a modulation frequency of 48 kc/s is shown. The first of these frequency bands is assumed to be well above the highest Doppler frequency of interest and will be used primarily for initial detection. The lower modulation frequency bands will only be used in conjunction with the "internal" F.M.C.W. mode. The reason for this is explained in the section on receivers.

The two frequencies ω_1 and ω_2 are also used to derive two local oscillators for the receiver which is divided into two parts. An upper sideband and a

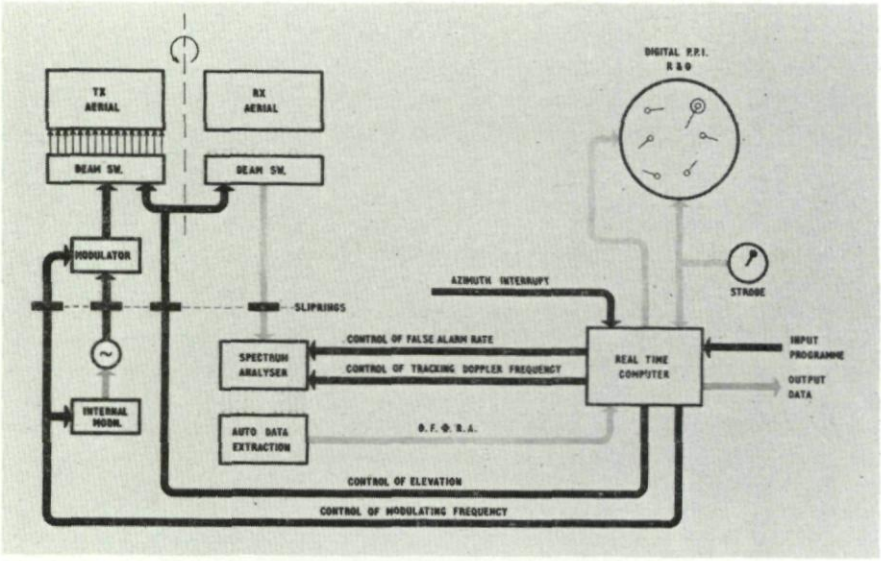


FIG. 1. *Automatic F.M.C.W. surveillance radar.*

lower sideband can therefore be selected, the order of the sideband being equal to $(\omega_1 - \omega_2)/2p$ where p is the final modulation frequency applied.

2.2. *Aerial System*

The characteristics of the aerial system under development are shown in more detail in Fig. 2. The aerial covers a total elevation of 0–60° and is divided into twentythree beams each covering roughly 3°. The azimuth beam is about 2° in width so that at the designed maximum rotation speed the beam will sweep through about 200 beamwidths per second. The minimum possible receiver bandwidth or resolvable Doppler band is thus of the order of 200 c/s, corresponding to a radial velocity of 3 m/s. (at *X*-band). At lower aerial rotation speeds this figure will be correspondingly lower.

Under normal conditions of operation all the available transmitter power will be switched into one elevation beam at a time and thus only one receiver will be required for all elevation angles. Under this condition the greatest detectability per scan is obtained with the minimum of equipment and the longest time is taken to scan a given volume of space. The elevation beam is

selected by means of the beam switch shown, under command of the real-time computer. The operational "programme" is discussed below. The transmitter and receiver aerials are similar and similar beam switching is employed in both transmitting and receiving directions. If desired, beams can be used in parallel, which of course necessitates multiple receivers.

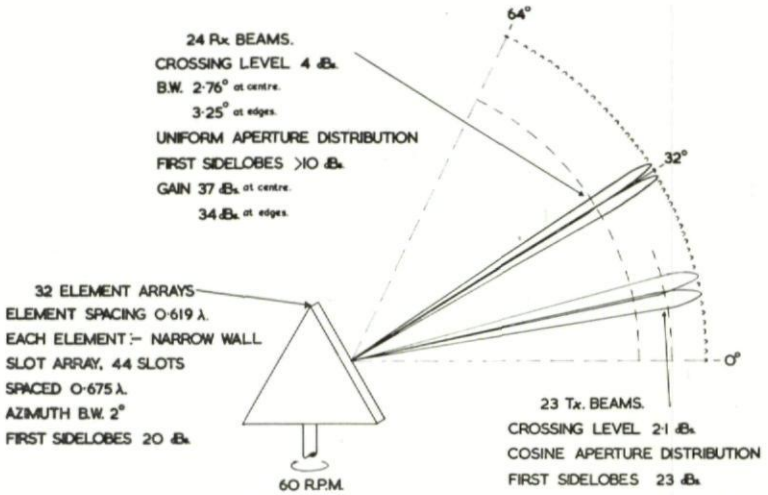


FIG. 2

2.3. *The Receiver*

The main features of the receiver are shown in Fig. 3, where the frequencies quoted are for the special case of a modulation frequency $f_m = 48$ kc/s. If ω is the transmitted carrier frequency, τ is the time delay to and from a target, ϕ is the modulation index, and ω_{IF} is the first intermediate frequency, the nth order sidebands can be derived[1–5] in the usual manner using Bessel Functions and are given by either

$$J_n(\phi_r) \cos [\omega_{IF}t + \omega\tau \pm np(t - \tau/2)]$$

or

$$J_n(\phi) \cos [\omega_{IF}t + \omega\tau \pm np(t - \tau)]$$

where $\phi_r = 2\phi \sin p\tau/2$ and $p = 2\pi f_m = (\omega_1 - \omega_2)/2n$, depending on whether one uses the F.M.C.W. or the unmodulated carrier as a reference. The amplitude terms $J_n(\phi_r)$ and $J_n(\phi)$ apply to a F.M. signal of unit amplitude and express the relative amplitude of sidebands.

By applying the two local oscillators derived from the modulator oscillators mentioned above, that is, by shifting the upper sideband downwards by a frequency $\omega_{IF} - \omega_2$ and the lower sideband downwards by a frequency $\omega_{IF} - \omega_2$, the sidebands are given by either

$$J_n(\phi_r) \cos [(\omega_1 + \omega_2)t/2 + \omega\tau \pm (\omega_1 - \omega_2)\tau/2]$$

or

$$J_n(\phi) \cos [(\omega_1 + \omega_2)\, t/2 + \omega\tau \pm (\omega_1 - \omega_2)\tau]$$

The signals are thus brought to a common frequency $(\omega_1 + \omega_2)/2 \approx$ 2.8 Mc/s and can each be passed through identical filters centred on frequencies $(\omega_1 + \omega_2)/2 + \omega\, d\tau/dt$, where $\omega\, d\tau/dt$ is the Doppler frequency. Similar pairs of filters are required to cover a given detection band (e.g. from +2 kc/s to +20 kc/s if initial detection of very slow or receding targets

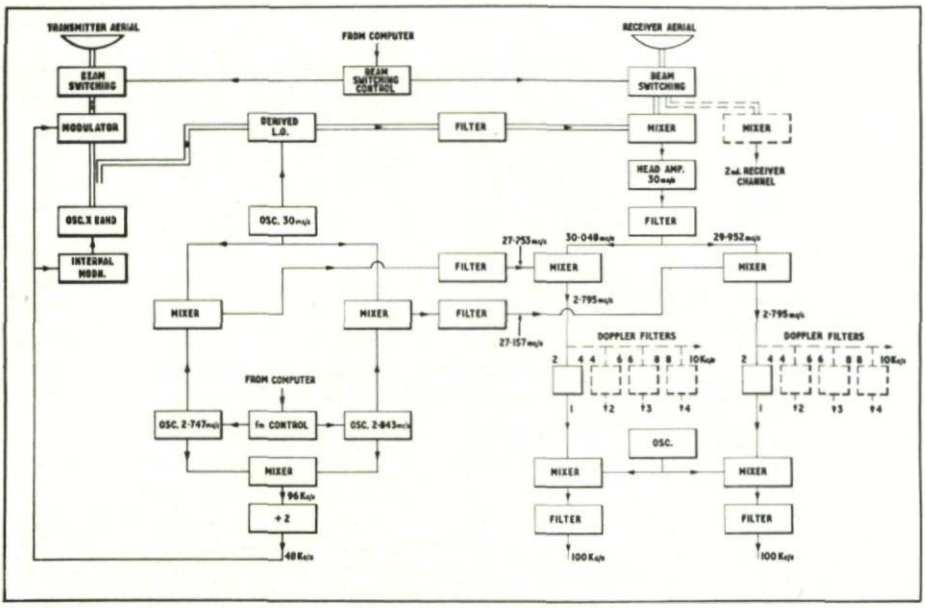


FIG. 3. *Two frequency F.M.C.W. radar.*

is not required). A separate, though closely similar receiver will be used for tracking. In this case the 30 Mc/s reference oscillator, ω_{IF}, is replaced by two separate oscillators of similar frequency. One of these is of fixed frequency and is used to derive the microwave local oscillator. The other consists of a set of crystal oscillators, one of which can be selected by the computer and is used to create the two 27 Mc/s reference oscillators. For complete surveillance this set must cover negative as well as positive Doppler frequencies.

The phase difference between the upper and lower sidebands is clearly $(\omega_1 - \omega_2)\tau/2$ in the F.M.C.W. case and $(\omega_1 - \omega_2)\tau$ in the two-frequency case. Using known frequencies ω_1 and ω_2, (and allowing for fixed phase shifts in the receiver) the range can be computed, but the result is at first ambiguous. It is proposed to use $\omega_1 - \omega_2 = 85$ kc/s for detection purposes and $\omega_1 - \omega_2 = 86$ kc/s for second paints, thus two paints are required to obtain a sufficiently unambiguous measurement, (namely to 300 km in the F.M.C.W. case or 150 km in the two-frequency case). Thereafter, if further accuracy is required, successive measurements can be made at frequencies of 84, 88, 80, 96, 64 kc/s, i.e. using successive frequency differences of 1, 2, 4, 8, 16, 32 kc/s. The method is very similar in principle to that employed in

the Tellurometer (see reference 5, for example) except that each measurement requires extrapolation since target motion between paints must be accounted for. It is expected that the Doppler frequency will be sufficiently reliable for this purpose, since quite large errors due to target modulation can be tolerated. If the Doppler frequency is not reliable enough an alternative process will be employed involving extra measurements. Assuming reliable Doppler frequencies the range error is reduced by one half after each measurement. After seven measurements the error is sufficiently small to remove all ambiguity and all further paints can be made using $\omega_1 - \omega_2 = 96$ kc/s. Thus ultimate range accuracy is a few percent of 3 km, or about 200 metres, depending on phase errors (using F.M.C.W. and the above-mentioned band of modulation frequencies).

It should be noted that the phase is independent of the order n of the sideband and that to select a given sideband pair only, a switching of the modulator divider to a value $2n$ is required. Using F.M.C.W., the relative amplitude of the sideband is $J_n(\phi_r) = J_n(2\phi \sin p\tau/2) = J_n[2\phi \sin(\omega_1 - \omega_2)\tau/4n]$ so that the distance between signal "holes" increases with n. The ambiguity diagram thus changes from a low Doppler-high range ambiguity type to a high Doppler-low range ambiguity type as information about the position builds up. By maintaining a constant deviation, i.e. an index which increases with n, and by computing a suitable modulation frequency, i.e. so that the expected signal frequency always appears between clutter lines in Doppler frequency and near an anti-node in amplitude, improved range resolution can be achieved with negligible loss of S/N ratio. In particular, rejection of very short range clutter can be obtained.[1-5]

Initial experiments using a 10 watt transmitter have suggested that the advantages of the F.M.C.W. mode, including the ease with which it can be applied, outweigh the disadvantages of the range "holes." The latter are inconvenient primarily in initial detection, where there is a roughly 20% chance of losing more than 3 db signal power on any one scan. If it is possible to "fill up" the holes either by making use of target motion, or by repeating scans using different modulation frequencies, the two-frequency mode can be dispensed with.

The receiver is essentially a spectrum analyser with parallel outputs. Ideally, to give optimum resolution, the number of pairs of paralled outputs is given by the total detection band divided by the minimum possible bandwidth and thus a very large number of channels would be required for detection of incoming targets only. However, a great economy in equipment can be obtained since optimum target resolution is unnecessary for initial detection and target detectability is not greatly affected by pre-detector bandwidth. For these reasons it is proposed to use 2 kc/s bandwidth filters, reducing by a factor of ten the number of parallel channels required.

Having obtained an initial detection, the computer estimates as closely as possible the time when the next paint is due and, at the appropriate moment, switches the elevation beam to the predicted elevation and sets a small bank of optimum bandwidth filters to the predicted Doppler frequency. As described above, this bank is capable of being positioned at any point in the

Doppler band and must cover a band of frequencies at least as wide as the uncertainty in prediction. After the first paint, azimuth and Doppler rates are known very roughly from initial assumptions about the target; thus the uncertainty may be, for example, 3 beamwidths in azimuth and 8 (fine) channels in Doppler. In this case, the false alarm probability per channel per azimuth beamwidth can be made somewhat greater than for initial detections as there are only $8 \times 3 = 24$ possible outputs. After the second paint the first differentials can be calculated and, since second differential effects are small, this leads to a smaller uncertainty—perhaps of the order of one beamwidth in azimuth and one channel in Doppler. The false alarm probability per channel per azimuth beamwidth can thus be greatly increased so that the probability of detecting third and subsequent paints approaches unity. Thus tracking through fades, for example, can be made more secure without much effect on the input data rate to the computer. The latter is largely governed by the first paint, false alarm rate.

It follows that the amplitude selection level, which determines the level of signal accepted and the false alarm rate, together with the Doppler bank setting and the number of channels required, must be controlled by the computer. By the use of coarse/fine Doppler banks in this way a very wide spectrum of radial velocities can be analysed economically, without wasting signal energy, initially into 30 m/s bands and subsequently into 3 m/s bands.

2.4. *The Automatic Data Extraction Equipment*

The combined outputs of the receiver, or spectrum analyser, thus consist of two sets of outputs, the first set consisting of perhaps ten detection outputs and the second of about eight tracking outputs. As shown in Fig. 1, these outputs are fed in parallel into the Automatic Data Extraction equipment whose purpose is to extract the required data from the inputs and to convert them into digital form. The data required are the co-ordinates and amplitude of each signal return. There are four co-ordinates viz: the azimuth angle θ, the Doppler frequency F, the elevation angle ϕ and the range R. The signal amplitude is also required, though not to a high order of accuracy, to help in associating co-ordinates. Theoretical RMS errors due to noise are $(1/2\pi)\sqrt{N/2S}$ measured in fractions of a beamwidth or of maximum unambiguous range, as the case may be, where N/S is the final noise/signal power ratio. Present indications are that, unless special precautions are taken, instrumental errors will be at least 5%, and so will have a major effect on the "bit length" required. At present the word lengths proposed are: eleven bits for θ, ϕ, and F and eight bits for R.

The auto extraction equipment can be divided into two parts: (1) analogue/digital conversion equipment on each channel and (2) a control unit which converts the primitive digits available from any channel into binary numbers representing θ, ϕ, R, F, and A. Referring to the receiver, the selected upper and lower sidebands are detected, combined, and smoothed in low pass filters. These filters are matched to the burst time so that on average only one signal or noise peak is obtained per beamwidth. To suit the extraction equipment the wide range of signal amplitudes possible is compressed,

prior to detection, to a much smaller amplitude range. Two digital outputs are obtained, one marking the instant in time when the energy in the combined channel is at a peak, and the other enabling the signal amplitude itself to be measured. The former determines the only instants in time (azimuth) at which the co-ordinates may be measured.

A separate process is required to obtain range. Each pair of sidebands is passed into phase comparison circuits whose outputs are linearly proportional to phase difference. This phase varies very slowly with time, at a rate less than 0.3 c/s or 0.5° during the signal burst. Two alternative outputs are required, since at the instant of measurement the phase may be changing from zero to 360°. These outputs differ by 180° and thus a further digital output is required to "advise" the control unit which of the two phase outputs to use. The control unit makes an appropriate correction to the most significant digit.

Further outputs are required to provide frequency measurement on the detection channels (i.e. measurement of Doppler frequency to a fraction of a channel width) and elevation measurement (to a fraction of a beamwidth). In addition, outputs of a logical nature are required to reduce the number of false detections due to imperfect equipment. These include digital outputs, which indicate that the output frequency of a channel is truly of that channel rather than "breakthrough" from an adjacent channel, and an additional detection output measuring the phase coherence of the ranging output, which eliminates certain spurious signals caused by an imperfect transmitter or receiver. These and other logical outputs are AND gated with peak pulse and amplitude comparator output to provide a "detection" output.

The Control Unit extraction programme is normally restarted on receipt of an azimuth pulse from a counter fed at about ten pulses per azimuth beamwidth turned; and its speed is such that about two extractions can be performed per azimuth pulse. This peak rate, which corresponds to about 20 per beamwidth or 4000 per aerial revolution, is some 50–100 times faster than the computer could accept, except for short periods. The mean rate is reduced to a reasonable level by taking only the larger signals from each channel. This level setting will be under the control of the computer as described in the previous section.

The control unit programme begins with an amplitude selection sub-programme. The detection AND-gate outputs from all channels are examined simultaneously for a detection, the reference amplitude being progressively reduced from high to minimum level. In most cases, no detections will be observed, for the reason explained above, but occasionally the signal will be large enough. At this point the amplitude level ceases to fall and the amplitude and azimuth of the signal can be transferred to the computer. The channel responsible for the detection is then found, whereupon the range, Doppler frequency and elevation can be measured. The computer programme is interrupted and the data transferred. If at the same amplitude, another detection is found in another channel, then this also is measured. Otherwise the amplitude selection sub-programme continues if time

is available. At some point in time the next azimuth pulse will arrive which resets the control unit and restarts the main programme.

Automatic radar control is initiated by a signal from a counter. This counter contains the current value of $\Delta\theta$, which is the angular difference between the current aerial azimuth and the azimuth angle at which some change in the radar properties is required. The quantity $\Delta\theta$, originally calculated by the computer and transferred to the counter, is fed with azimuth pulses until it contains zero. When this occurs, the auto-extraction programme is discontinued and the computer changes the modulation frequency, the elevation beam, the tracking Doppler frequency, the false alarm rate (or minimum detection level) and switches on or off channels as required.

2.5. *The Real-Time Computer*

This is a general purpose computer normally carrying out a predetermined programme (see Fig. 1: "Input programme") which can be nevertheless interrupted at the end of a current instruction by one of a number of peripheral devices. Each device is allocated a fixed priority to avoid any problem arising from simultaneous calls from different devices. When a device gains control over the computer, either a data transfer, or a programme jump, or both, takes place.

The inputs consist of main and interrupt programmes, which are normally read in advance; data inputs and calls from the auto-extraction equipment; and manual interrupts (i.e. from keys or the "strobe"). The strobe is used in conjunction with the digital P.P.I. (see below).

The outputs may be divided into control outputs and data outputs. The former have been already described and go to the elevation switch, the modulator, the tracking Doppler bank and the auto-extraction equipment. The data outputs would, in a developed system, depend greatly on the application of the radar system. For example, in a G.W. system the computer might also be required to control a tracking radar. In an A.T.C. application, collision warning may be required. In the system presently under investigation, it is presumed that the objective is to obtain tracks of moving targets, which is the basic requirement for all applications. It is thus essential to output track data together with any other data required for the experimental investigation of the system. Some of these data can be punched on tape in the usual way, but in a real-time system an enormous quantity of transitory data is created. To cope with this it is proposed to use a digital P.P.I. which can display tracks in two dimensions at a relatively fast rate (up to 200 tracks per second).

In Fig. 1. typical tracks are shown in (R, θ), where "present" (R, θ) is indicated by a characteristic small circle with a line drawn to "future" (R, θ) as predicted by the computer. Around one of the tracks is shown a larger circle, which has been created by inputs from the manually operated strobe. These data from the strobe can also be transferred to the computer and any desired operation with regard to that particular track may be

carried out (e.g. outputting all available data; deleting from the store; causing the auto-extraction equipment to eliminate the channel on the next revolution of the aerial etc.).

3. PROGRAMMING AND OPERATIONAL CONTROL

It is not proposed to deal with this subject in any but a superficial way, partly because much of the detailed programming has not been worked out and partly because these details will vary with experimental requirements. The programmes can be conveniently divided into Main and Interrupt programmes, although in practice the distinction may be greatly blurred.

3.1. *Main Programmes*

(i) *Initial Detection.* In almost all cases it can be assumed that one has, a priori, information concerning the likely directions and radial velocities of targets. Thus, for example, we would normally expect to detect aircraft initially at relatively low angles of elevation and at relatively high radial velocities. In addition, one would expect, at least in a land based system, that the contours of the territory surrounding the radar would be such that in some directions targets would be detected only at higher angles. However, it is not a difficult matter to store such geographical details, to a sufficient order of accuracy, in the computer. It follows that the initial detection programme would consist of, perhaps, one or two variations in elevation angle during each rotation of the beam, (i.e. to follow the contours) coupled with a change in elevation angle every rotation of the aerial. For detection of aircraft, the overall scan in elevation might be less than three or four beam-widths.

A list of numbers appropriate to the circumstances would thus be kept in the store. These numbers would refer to the azimuth angle at which some change is required, to the required elevation angle, and to the detection Doppler band to be used. The modulation frequency would be set to the first in sequence. To stabilise the system it is essential that this programme be of rather low priority.

(ii) *Track Prediction.* Any outputs from the auto-extraction equipment would be placed on a list, together with pointers to the appropriate track if obtained via the tracking bank. In the absence of interruptions and at a rate governed only by the internal speed of the computer, new co-ordinates would be compared with predicted values of θ, ϕ, and F so as to associate them with existing tracks. Any associations so formed would be used to update the track, i.e. predict the new co-ordinates and estabilish the "box" size. The "box" is the uncertainty in θ, ϕ, and F (and R if range has been resolved) and would be diminished as the track is built up. In the case of new tracks, the next stage of range resolution would be performed, but, in the case of first paints (i.e. arriving via the detection bank), only the box size could be calculated.

At this stage an established track could be fed either to the digital P.P.I. or to the punch. A decision would than be made as to the number of revolutions of the aerial which could be allowed before the next "paint" was required. Established and range-resolved tracks could be updated internally (i.e. without recourse to new information) for several revolutions, but new tracks would need new data as soon as possible. A priority system is therefore required to avoid losing important tracks. When the maximum time that can be allowed to elapse before a new paint is required has been determined, the new box size would be calculated and the data replaced in a list of increasing $\theta - \delta\theta$, where θ is the predicted next azimuth and $2{\cdot}\delta\theta$ is the box size in azimuth. New store addresses would then be allocated for the future data.

(iii) *Test Routines.* Assuming the main programme was well in advance of time, and subject, of course, to the usual interrupt programmes, spare time would be used to give warning of any malfunction of the system.

3.2. *Interrupt Programmes*

(i) *Automatic Radar Control.* This programme is started by the "time to go" register, i.e. when $\Delta\theta = 0$. The main programme is interrupted and new values of ϕ, F (tracking), modulation frequency, false alarm rate, etc. are inserted. The next value of $\Delta\theta$ is also inserted and the required adjustments in the queue are made. A priority system will be required here to take account of simultaneous requirements on the radar; any missed paints must be noted and attempted on the next revolution of the aerial.

(ii) *Automatic Extraction Outputs.* This programme is relatively straightforward and involves merely a transfer of the quantity (θ, ϕ, F, R, A) to a previously named store and the preparation of the next address. If the output is from the tracking bank, the initial association in θ, ϕ, and F has already been made, since the computer itself caused the output; thus the linking address is already in position. In some cases more than one output may be linked with a given track, but these will tend to be eliminated in the main programme. In those cases where multiple associations cannot be immediately eliminated the track must divide. It is essential, therefore, that the probability of track division be not so high that a net growth of tracks is obtained. If the output originates from the initial detection bank, of course, no linking address is necessary. The fact that the output is from the detection bank rather than from the tracking bank can be inferred from the channel number part of F, so that it will be possible to operate both banks simultaneously if required.

(iii) *Strobe Call.* In this case the (R, θ) digits are transferred to the computer. The track list is searched through until the correct association is made, at which point any predetermined programme may be carried out on the particular track. In experimental work this would consist of hard copy of all data relevant to the track (via the computer punch). Other kinds of manual interference with the normal operation of the system can, of course, be achieved via the computer control console.

4. CONCLUSIONS

The experimental work has reached an intermediate stage in which the various parts of the system have still to be connected together. Thus an conclusions drawn must be tentative. However it seems clear that the connection of a digital computer to a radar must represent more than a replacement of a human operator if the full advantages are to be reaped. It has been shown that, if the flexibility of the computer is matched by a flexible aerial and waveform, a single radar can have a better performance in detection and tracking than a conventional two dimensional radar and at the same time handle many tracks both of long range and of low level targets. It has also been shown that the F.M.C.W. waveform can be used in such a way that the targets are detected firstly as moving objects of virtually unknown range. One rotation of the aerial later the range can be determined to a rough order of accuracy and on subsequent rotations, as the track information builds up, to a progressively greater degree of precision. Final range accuracy is limited only by the modulation frequencies available. As the range becomes better known, higher order sidebands can be used to improve the low speed performance of the radar. This should assist in the tracking of crossing targets.

5. ACKNOWLEDGEMENTS

I am indebted to the Director of the Royal Radar Establishment, Ministry of Aviation, for permission to publish this paper. I particularly wish to thank Mr. A. Smart, who was responsible to a large extent for the original conception behind this work, for his continued interest and encouragement. I also wish to thank Dr. J. R. Prior, who is responsible for the modulation sequence, for many useful discussions. Acknowledgement is due to Mr. J. Wood, who designed the auto-extraction equipment, to Mr. A. J. Mould for his work on the radar and to several others for their assistance. Ferranti Ltd., Edinburgh, are currently responsible for the design and construction of the multi-beam aerial.

REFERENCES

1. SAUNDERS, W. K., *Trans. I.R.E.*, **AP-4,** 1956.
2. GLEGG, K. C. M. *DRTE/EL Report.* **5040-1,** Defence Research Board of Canada, 1956.
3. BROWN, R. K., *Proc. East Coast Conf. Aero. and Nav. Electron.*, 1958.
4. FRIED, W. R., *Trans. IEEE.* **AS-1**, No. 2, 1963.
5. SKOLNIK, M. I., *Introduction to Radar Systems*, **106,** McGraw-Hill, 1962.

CHAPTER 16

DOPPLER SYSTEM FOR MEASURING THE FLIGHT PERFORMANCE OF SMALL ROCKETS AND LIGHT AMMUNITION

C. O. Lund
Copenhagen, Denmark

A Doppler System for the measurement of the velocity of missiles and ammunition is described.

The basis of the design was a conventional type of Doppler radar with a CW magnetron transmitter and a receiver with IF amplifier.

Due to stringent performance specifications, various improvements were introduced:

(a) *The magnetron was operated in a mode with low oscillator noise.*
(b) *The magnetron was stabilized with a high-Q cavity.*
(c) *A special antenna system with low RF leakage between the transmitting and receiving antenna was developed.*
(d) *Special non-tracking-filters were used in the Doppler signal path.*

The Doppler radar developed and produced has at present a range of approximately 1600 *meters for* 5″ *rockets, but various improvements have been contemplated and some have been tested. A range improvement by a factor of* 4 *seems feasible.*

1. INTRODUCTION

Doppler radars for measurement of missile velocity have been known for some time. In this paper will be described a system built to specifications which cannot be met by the simple type with direct detection.

The radar was originally developed and produced for Raufoss Ammunisjonsfabrikker, Norway by Terma Electronic Industry A/S, Aarhus, Denmark.

Also participating was The Microwave Laboratory under the Danish Academy of Technical Sciences, Copenhagen, which handled certain development works,—mainly concerning the antenna system and the low-noise transmitter—for Terma.

The author acted as inspector/consultant for Raufoss and followed all phases of development and construction.

2. SYSTEM CONSIDERATIONS

2.1. *Primary Specifications*

(a) Minimum target area $\sigma = 0.01$ m²
(this corresponds to a missile diameter of app. 5″)

(b) Maximum range $d_{\max} = 1000$ meters.
(c) Velocity range (max) $v = 0$–600 m/s.
(d) Minimum antenna beamwidths, both horizontal and vertical $\theta = 16° (\pm 8°)$.
(e) Accuracy of velocity recording $\pm 3\%$.

Further the equipment has to operate in a most noisy environment, the maximum noise levels were not part of this specification, but every reasonable precaution against microphony in the equipment was taken during the initial design.

Various factors are fixed by the primary specifications.

The antenna gain is limited upwards by the specifications on beamwidth. The corresponding gain is then maximum $G_A = 21$ db independent of operating frequency.

The bandwidth is dependent on operating frequency and the main decision to be made concerns frequency and transmitter power.

2.2. *Choice of Operating Frequency*

The radar equation:

$$P_R = P_T \frac{c^2}{(4\pi)^3} \frac{G_A{}^2 \sigma}{f_S{}^2 d^2}$$

where

P_R: Received Power.
P_T: Transmitter output power.
c : Velocity of light.
G_A: Antenna gain (assume the same gain for transmitting and receiving antenna).
σ : Target reflection area.
f_S : Operating frequency.
d : Range.

The received signal power P_R will have to compete with various sources of noise, primarily the receiver noise (assuming a superheterodyne receiver with noise factor F):

$$P_{NR} = kT_0BF,$$

where the bandwidth $B = \alpha \cdot f_D$ is determined by the signal bandwidth, again proportional to the Doppler frequency

$$f_{D\max} = \frac{2V_{\max}}{c} \cdot f_S$$

where $V_{\max}$ is the maximum target velocity.

If we combine these equations we get an expression for the signal-to-noise ratio as the ratio between signal power P_R and receiver noise P_{NR}:

$$\frac{P_R}{P_{RN}} = \left[\frac{c^3}{(4\pi)^3 \cdot 2 \cdot kT_0F}\right]\left[\frac{G_A{}^2 \cdot \sigma}{d^4}\right]\left[\frac{P_T}{\sigma f_S{}^3}\right]$$

The first bracket is more or less fixed by laws of nature and practical considerations and the second bracket is determined by the specifications.

The system design has then to be based on the third bracket where we see that the required transmitter power increases proportional to the third power of the operating frequency. For example, 10 Gc/s requires 30 times more transmitter output power than 3 Gc/s.

A fairly low frequency should then be used, mainly limited by the practical size of antennas and determined especially by suitable transmitter tubes and available frequency bands.

These considerations led to an operating frequency of $f_S = 2.45$ Gc/s based on a magnetron transmitting tube.

2.3. *The Basic System*

Figure 1 shows a simplified block diagram for the radar.

The signal from the transmitter is attenuated by L_T in the transmission path, this includes the antenna gains $G_A = 21$ db, the maximum distance $d_{\max} = 1000$ m and target reflection area $\sigma = 10^{-2}\,\mathrm{m}^2$.

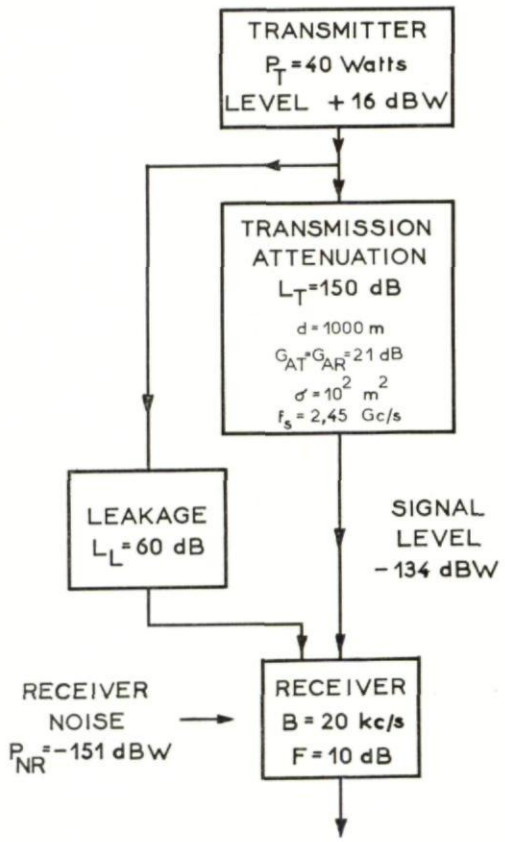


FIG. 1. *System parameters.*

The corresponding attenuation is $L_T = 150$ db.

The signal level at receiver P_R is then computed to -134 dbW for 40 W transmitter output.

For $fs = 2.45$ Gc/s and $V_{\max} = 600$ m/s we have the maximum Doppler frequency $f_{D\max} = 9.8$ kc/s and an estimated bandwidth (modified later) of $B = 20$ kc/s.

The estimated level of receiver noise (noise factor $F = 10$ db) is then $P_{NR} = 151$ db.

The (ideal) signal-to-noise ratio is then 17 db, this was considered adequate for further planning.

2.4. *Leakage*

One of the most troublesome factors in CW radar systems is the leakage noise from transmitting to receiving antenna.

An estimate of the noise level is based on a leakage attenuation of 60 db, if the contribution of transmitter noise should be less than receiver noise, then P_{NT}, the transmitter noise in the Doppler-frequency band, should be less than -91 dbW.

This problem was attacked both by finding an advantageous operating region for the transmitter magnetron and also by choosing an antenna type for minimum leakage.

Another factor is the leaked CW power, the level at receiver is estimated to -44 dbW, so, in the IF amplifier, the modulation ratio for the Doppler signal is -90 db; the IF amplifier has to be designed for this.

2.5. *Clutter Modulation*

Clutter, due to reflections from nearby moving objects such as vegetation and water waves influences the system.

As a self-oscillating transmitter is used, back reflection will cause modulation of the transmitter output, even if this is fairly small, it will have a smearing effect on the Doppler signal and should be limited.

A frequency stabilizing circuit with a high Q resonator gives a high degree of stabilization (stabilization factor 75).

Clutter interfering in the signal path can be limited by proper choice of filters for Doppler signals and by proper siting of the radar.

3. DESCRIPTION OF THE EQUIPMENT

A diagram of the main units is shown on Fig. 2. The main features of the equipment are:

3.1. *Transmitter*

This is based on a Philips CW magnetron type 7090, operated at reduced voltage and nominal current by reducing the magnetic field.

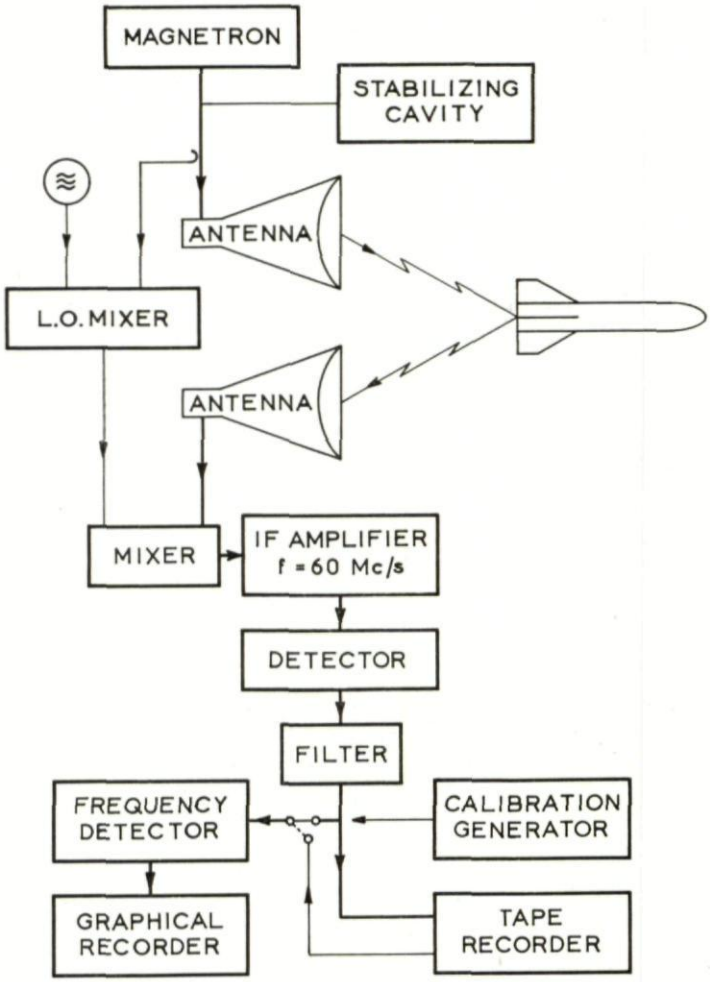


FIG. 2. *System block diagram.*

The nominal output power for this magnetron is 100 W, at that operating point, the noise level in 23 kc/s bandwidth close to the carrier is quite high (app. −40 dbW). When the anode voltage is reduced to 800 V and the anode current kept constant by reducing the magnetic field, the noise level drops suddenly to a level around −80 to −90 dbW. At this operating point, the output power is 60 W, leaving some margin for loss in the stabilizing circuits for 40 W to the antenna.

The stabilizing circuits were designed, according to standard practice, with a brass cavity; no attempts were made to make temperature compensation, the cabinet for the transmitter etc. has temperature regulation and the specifications for accuracy in velocity measurements are not too strict.

3.2. *Antennas*

The primary specifications and the choice of operating frequency decide the aperture area corresponding to 21 db gain.

A rather unconventional type of antenna, the diagonally excited pyramidal horn with dielectric lens was chosen because:

(a) Both the horn and the lens can be made rigid with maximum protection against modulation due to mechanical vibrations. No difficulties have been encountered with the antennas as built.

(b) The diagonal excitation gives a higher attenuation against leakage from one antenna to the other than most other antenna types. The attenuation measured is higher than 75 db.

FIG. 3. *Transmitter-receiver unit.*

The horns as mounted in the transmitter/receiver unit are shown in Fig. 3, the aperture side is 400 mm and the horn length 600 mm, 5 mm aluminium alloy is used, the lenses are made from Lucite, each weighs 14 kg.

3.3. *RF Transmission Lines*

Except for short waveguide sections at the horn terminals, co-axial lines are used in the transmission circuits. This type of line was used rather than waveguides in order to minimize the effect of vibration. It can be shown that the impedance of the waveguide varies directly with the relative deflection due to vibration, whereas for co-axial line, impedance varies only by approximately the square root of the relative deflections. It is believed that with proper design, co-axial lines have more protection against microphonic effects than do waveguides.

3.4. *Mixers*

The local oscillator signal, 60 Mc/s, is derived from the transmitter by a conventional mixer, the second mixer is the signal mixer, operating with about 0.5 mA crystal current.

There are adjustable filters in both channels, these filters are built in a common block and the mixers (commercial co-axial types) are mounted on this.

3.5. *IF Amplifier and Detector*

The IF amplifier is of special design, it has to operate with a very low modulation index due to the leaked transmitter signal.

In order to increase the modulation index, the amplifier is operated in class C and the minimum modulation index is increased from −90 db to about −70 db.

The output of the IF amplifier is detected directly, the leaked transmitter signal can be used as a carrier.

The IF amplifier uses ordinary electron tubes, in order to avoid microphony,[2] a transistor type IF amplifier would be desirable. However, no disastrous microphony has been detected in the amplifier. The mixers are the only units that gave some difficulties with microphony, these difficulties were traced to a detector contact, but after redesign no difficulties were noticed neither with 5″ rockets nor 105 mm guns fired within 10 meters of the transmitter receiver unit, nor with submachine guns fired with the muzzle 10 cm above the top of the unit.

3.6. *Filter*

The foregoing sections of the radar do not have to be specially designed for any special range of target velocity. However, the filter will limit the noise bandwidth and has to be designed for each velocity range and target type.

As stated before, the radar was primarily designed for velocity measurement on rockets, launched from a position very close to the radar. It is a characteristic of rockets that they accelerate over the first portion of the range. Thus, when the rocket is close to the radar, the Doppler frequency is low and the signal is high, and the signal decreases and the Doppler frequency increases with distance.

This can be utilized to narrow down the band-width for the Doppler signal by using a band-pass filter, where the attenuation in the low frequency part decreases with approximately the sixth power of the frequency in the velocity range considered and increases fast above the maximum Doppler frequency.

The effective noise bandwidth in the Doppler channel will then be limited to approximately 2.5 kc/s compared to the 20 kc/s necessary, if the equipment were designed for all velocities from 0 to 600 meters per second at maximum distance. This improves the range performance by about 8 db.

3.7. *Data Handling Equipment*

Rocket launchings are costly and graphical recorders not too dependable, therefore, each measurement is recorded on a graphical recorder and in addition the Doppler signal is recorded directly using an ordinary tape recorder. The range of Doppler frequencies corresponds very well to the capability of a HiFi recorder.

The Doppler signal recorded on the tape can be played back to the frequency detector, also the tape recording can be used for further signal processing by tracking filters, etc.

The frequency detector is of the countertype, with a monostable multivibrator charging a condensor, the recorder measures the condensor voltage.

The pulsewidth of the multivibrator is chosen to correspond to a period of the highest frequency measured, it can be shown that this type of counter has

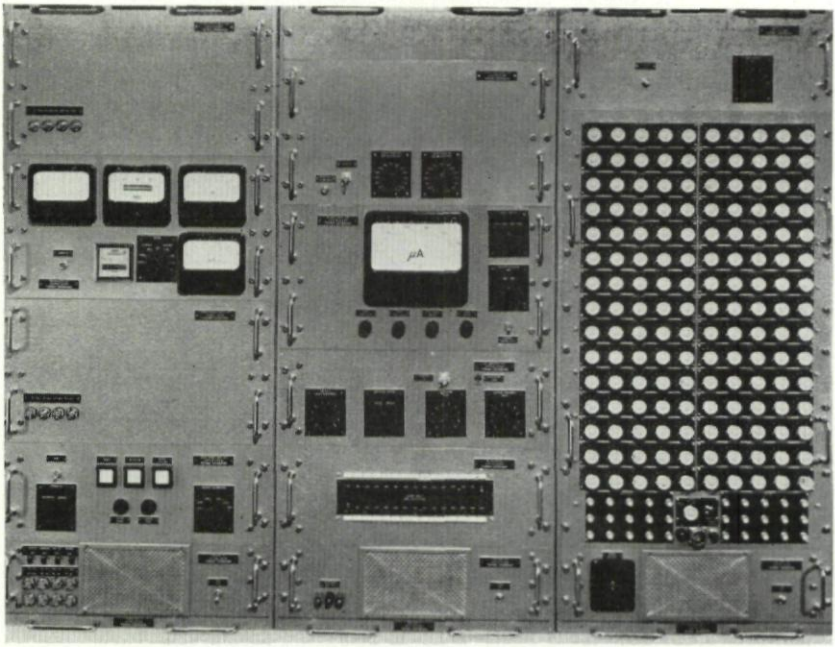


FIG. 4. *Data handling rack.*

higher immunity against noise of the higher frequencies than of the lower frequencies, this feature also improves the performance, so that the required signal-to-noise ratio at the highest velocity is about 2 db.

Figure 4 shows the rack for data handling equipment, calibration generator and timing equipment for firing. It is quite low so that it can be mounted in a small lorry.

3.8. *Calibration Facilities*

The velocity recording can be calibrated by known and fixed frequencies fed to the frequency meter. The calibration units are controlled by a separate timing circuit. This circuit will, when initiated by a firing contact, automatically start both graphical and magnetic tape recorders, make the calibration unit go through 5 velocity steps of calibration, disconnect the calibration unit and fire the missile and, after a reasonable length of time, stop the recorders again.

The calibration frequencies are derived as harmonics of a standard frequency chosen so that they correspond to reasonable values of velocity in meters per second, the velocity corresponding to a particular calibration step can be chosen by a set of switches on the calibration unit.

The RF output is checked by a simple power meter in the transmitter circuit and the transmitting frequency by the calibration of the tunable stabilization cavity. The performance of the receiver can in principle be checked by a sort of signal generator or modulated diode. If such equipment were to have a good accuracy, it would be rather complicated and it was decided not to include such facilities. Instead, at the start of each set of measurements a check on the performance is made by measuring the range performance with a caliber 7.62 mm projectile fired from an ordinary carbine. The realized range is about 80 m when the radar is functioning correctly.

This check has proved quite dependable and is easily performed by the launch crew.

4. PERFORMANCE

As described before the equipment was developed and produced on a crash basis with several parallel lines of development in order to meet the required

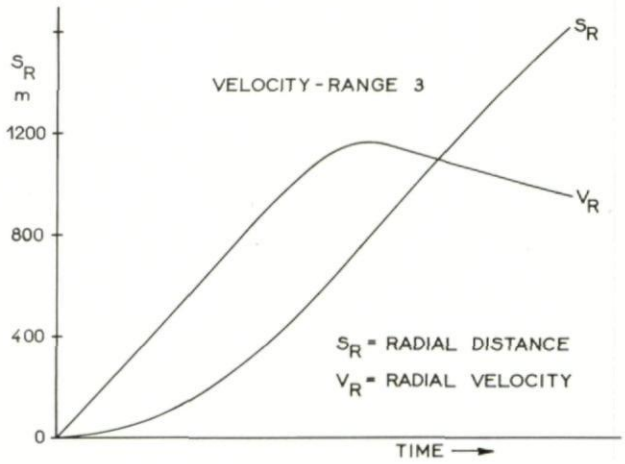


FIG. 5. *Missile distance and velocity.*

performance specifications. This resulted in the first set of firings in a recorded maximum range of 1600 m corresponding to RF performance margin of 8 db compared to specifications. In Fig. 5 is shown a redrawn recording, for security reasons I cannot show the time scale.

For the projectiles, caliber 7.62 mm was used for checkout of the equipment, the usual recorded range is 80 m. However, a magnetic tape recording of one of these tests was taken to a laboratory and analyzed with a Sonagraph.

In Fig. 6 is shown a picture where the main trace can be seen. The additional traces are due to intermodulation with the bias oscillator of the tape recorder and perhaps some harmonics.

On the original trace the velocity recording for this small bullet, with reflection area 1 mm,^{2} can be seen up to distances of 370 meters. Another interesting feature that can be seen in the original is that deceleration of the bullet is not uniform. This is due to the nutating movements of the bullet.

For automatic weapons, this Doppler radar is apparently quite useful. In Fig. 7 is shown a direct recording from a 9 mm submachine gun with nominal initial velocity of 400 m/s. The velocity of each separate bullet can be checked as well as the deceleration and the firing sequence. In this case

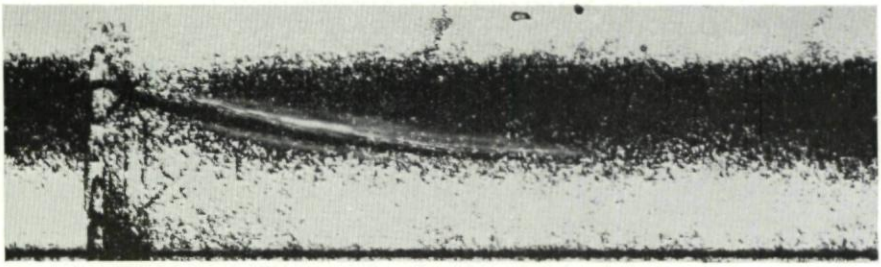


FIG. 6. *Velocity for* 7.62 *mm bullet.*

the nearest bullet will catch a Doppler signal so the recording is not range-limited.

Although the system is in accordance with specifications by a wide margin, improvements are possible. Automatic tracking filters with narrow bandwidth are a near possibility, in order to test this, one of the early recordings on magnetic tape was processed in an automatic signal tracking filter, re-recorded on magnetic tape, the graphical recording of this "cleaned" signal is shown on Fig. 8.

The maximum recorded range is 3.300 m. This is however not the maximum range, actually the original magnetic recording was stopped at the time of recorded maximum range. On the upper part is shown an example of the velocity-calibration described before, these calibrations seem a little noisy; this is due to rerecording made at least 3 times, some of them with 50 c/s power and some with 60 c/s power.

An automatic tracking filter can of course improve the operation of any Doppler radar, it is estimated that proper filters can improve the range capability at least by a factor of 4. As ranges of 1600 m on 5″ missiles have been achieved without a tracking filter, ranges in excess of 6 km can be visualized for this size of missile.

Because the recording facility at present is the limiting factor, the ultimate accuracy is not definitely known. With regard to frequency stability, the accuracy is probably around 0.2%, but this can be improved by proper calibration. For the present application, this accuracy is more than sufficient.

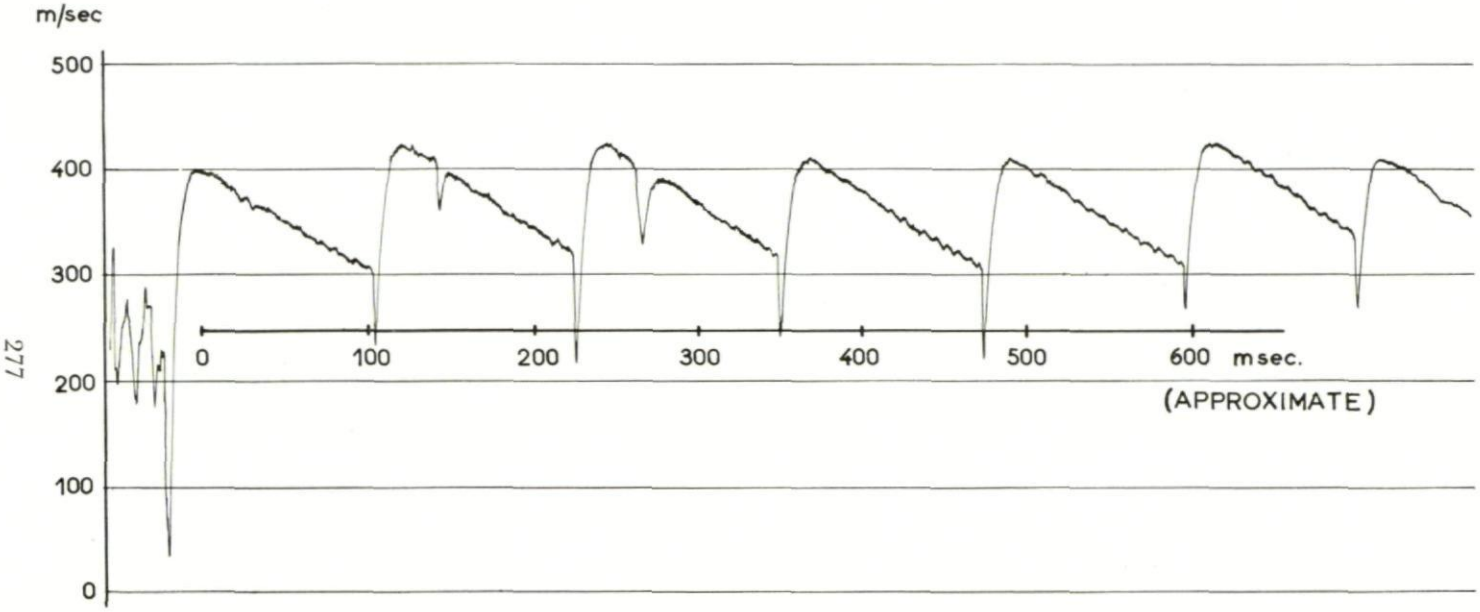


FIG. 7. *Velocity for sub-machine gun bullet.*

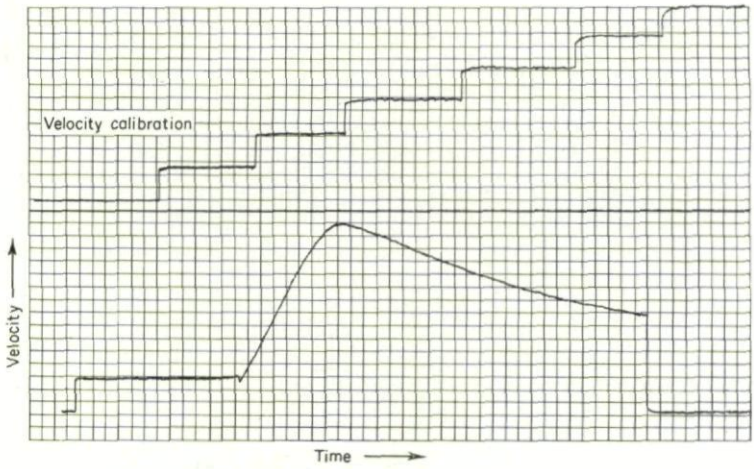


FIG. 8. *Missile velocity tracking filter analysis.*

5. RELIABILITY

The nature of the project has not made any formal reliability study and investigation possible, but the operational requirements were high considering the high costs of launch preparations and the distance of repair facilities.

The following precautions were taken in order to have high reliability:

(a) Sturdy chassis and racks.
(b) High requirements to workmanship in wiring, soldering and assembly.
(c) Components are highly derated and the majority of components and parts used were listed as type-tested by a qualified organisation.
(d) Electrolytic condensors were avoided—with one exception.

Until April 1964, two years after delivery, 500 operation hours have been logged, no large repair or adjustment has been necessary and in no case have radar breakdowns interfered with the launching program.

It must be mentioned that the radar has been transported in all several thousand kilometers on not too good mountain roads.

6. CONCLUSION

The design was started in August 1961 and the equipment was delivered at Raufoss early in March 1962, so, it is perhaps not necessary to say that the work was done in an atmosphere of high spirits and cooperation.

It was a rush-job, initially with a number of unknown factors and several possible roads for improvement. Several of these roads were followed simultaneously, resulting in a wide performance margin.

Improvements are of course possible, but so far, the described Doppler radar can measure the velocity of most missiles and projectiles over a fair proportion of their maximum range.

7. DISCUSSION

M. MORGAN: Can the speaker give me some indication of the accuracy that he is attaining in velocity and acceleration?

C. O. LUND: The original specification called for only $\pm 3\%$ accuracy in velocity. The accuracy was really limited by the graphical recorder to about 1% but just before I left I received the information that they were using a better recorder with which an accuracy of 0.2% is at least obtained under controlled conditions and there is no reason to say that it could not be achieved in ordinary operation.

A. R. DOMVILLE: Twelve centimeters is an excellent wavelength to choose for detecting raindrops. I take it you can only use your equipment in fine weather?

C. O. LUND: I cannot give a complete answer. I know that it has been used in snow storms but I agree with you that it would be vulnerable to rain, on the other hand we didn't take too much trouble in that line because this equipment was designed for rocket launching and testing which can only take place in fine weather. On the other hand—mosquitos are a problem.

M. MORGAN: If you use a tracking filter, one can usually manage to differentiate between the raindrops and the information from the target.

D. BOSMAN: We have made some very crude experiments and found that with a 30 mW output a range of about 400 meters was easily achieved on a walking person or a cyclist.

CHAPTER 17

INTEGRATED TRAJECTORY SYSTEM

Gus F. Bigelow
U.S. Army Electronics Research and Development Activity,
White Sands Missile Range, New Mexico, U.S.A.

A medium to high accuracy, multiple-target tracking system is being developed at White Sands Missile Range. Range and angle measuring equipment will employ continuous wave (CW) phase comparison measurements. Velocity measuring equipment will employ CW Doppler measurements.

1. INTRODUCTION

An Integrated Trajectory System (ITS) is being developed to instrument the White Sands Missile Range (WSMR) for measuring spatial positions, velocities, accelerations, and both scalar and vector miss distances for a variety of targets. When fully developed, ITS will provide these data in real time.

The important features of this system will be:

(1) full time, fast tracking in the hemisphere above the station,
(2) automatic acquisition when the tracking signal rises above the threshold,
(3) simultaneous tracking of up to five targets, where any combination of missiles, aircraft, drones, etc., may constitute the "targets" being tracked,
(4) range accuracy to less than ± 10 ft and direction cosine accuracy to ± 100 parts per million (ppm),
(5) range precision to ± 2 ft and direction cosine precision to ± 20 ppm, and
(6) reliable tracking at low initial and maintenance costs.

Three major subsystems will be used; these are:

(1) Distance Measuring Equipment (DME), in which at least three target stations are required to uniquely determine the spatial position of the target
(2) Angle Measuring Equipment (AME), in which at least two stations are required for the solution of the spatial position, and
(3) Doppler Measuring Equipment (Doppler ME), in which one transmitting station and at least three receiving stations are required. (Six or more receiving stations will actually be used.) The approximate locations of the AME and DME stations at WSMR are shown in Fig. 1. The Doppler ME, unlike the DME and the AME which are

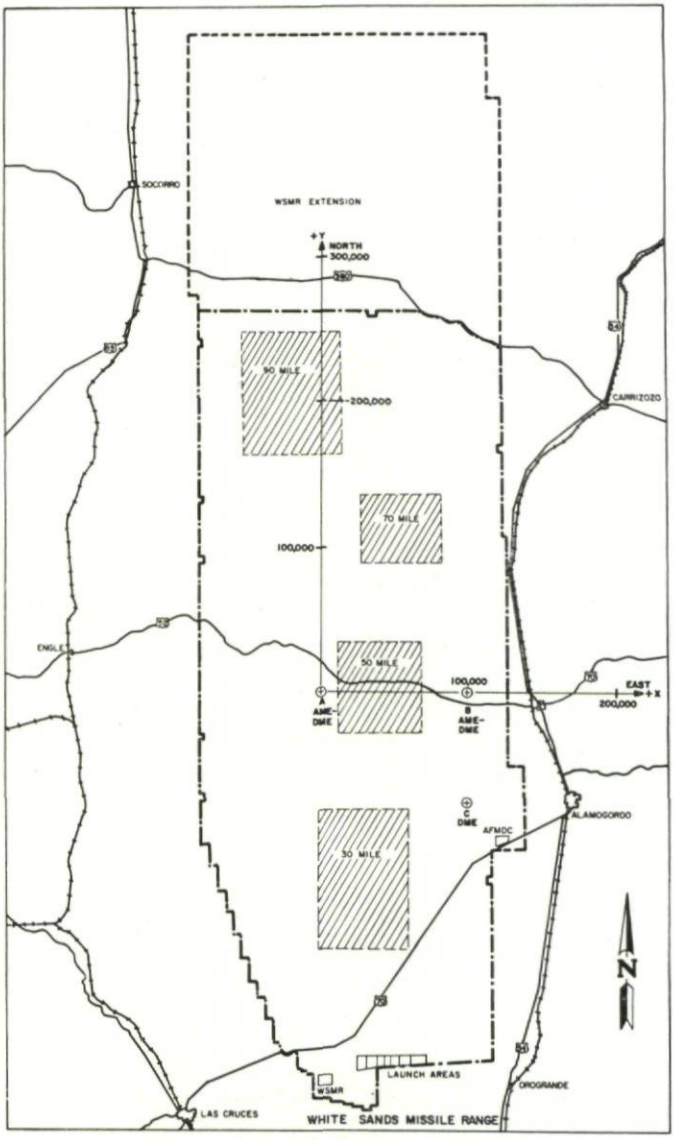


FIG. 1. *ITS station locations at WSMR.*

near-operational, is only now being designed, but it is planned that three of the Doppler ME stations will be integrated into the present DME sites and three will be located at other sites and be operated independently.

The DME is capable of tracking three targets simultaneously. Each target carries a transponder which each of three DME ground stations sequentially

interrogate. Each measurement results in an unambiguous range, and a spatial position is computed from the three resulting range measurements.

The AME is capable of tracking two targets simultaneously and requires only the reception of a signal from each target,—typically that of a telemetry transmitter. Two unambiguous direction cosines (per target) are derived at each of two stations, from which a unique spatial position is then calculated.

The Doppler ME is presently being designed to track one target only and will be nearly identical, in design, to the present DOVAP system.[14] (The system will be designed, however, for future expansion to multiple target

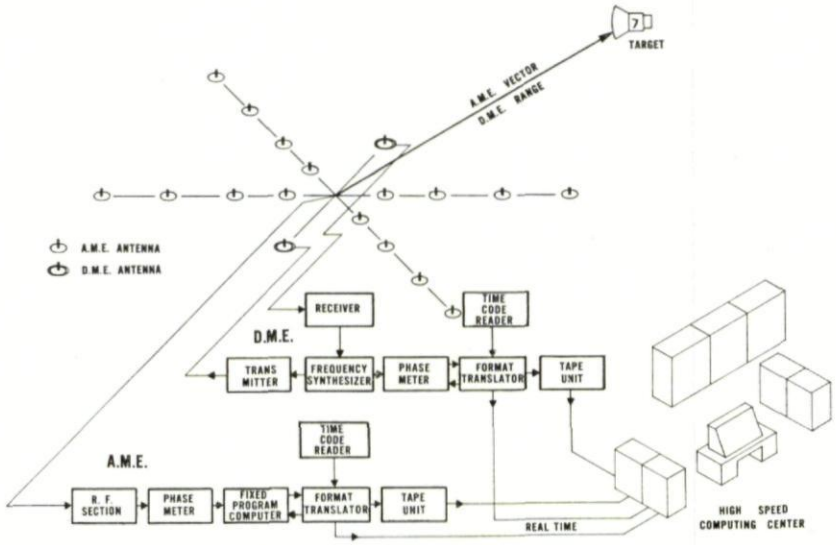


FIG. 2. *Integrated trajectory system AME/DME site.*

tracking.) It is to be compatible with the DME in that it will be able to operate on the DME transponder frequency. From range sum changes measured by the Doppler ME, spatial position, velocity, and acceleration will be calculated.

ITS may be operated in modes other than those just discussed, e.g., an AME/DME. This system configuration would measure two angles and a range and would therefore be analogous to a monostatic radar. Although mutual coupling between the AME and DME antennas presently precludes its immediate implementation, the AME/DME configuration is quite practical, and a workable system has been proposed. Fig. 2 is a block diagram of this proposed AME/DME.

1. PRINCIPLES OF OPERATION

A. *Phase Delay or Ranging (DME)*

The DME operates upon the principle of determining the one-way distance or range from a ground station to an airborne target by measuring

the phase delay of a CW signal returned from the target. Neglecting atmospheric effects, the propagation time of a phase front from a station to target and back is

$$\Delta t = 2R/c = 2R/\lambda f,$$

where R is the range, c is the vacuum speed of light, and λ and f are the wavelength and frequency, respectively, of the ranging signal. The phase delay which occurs during the transit time Δt is

$$\Delta\phi = 2\pi f \Delta t = 4\pi R/\lambda,$$

and the range is

$$R = \Delta\phi_m \lambda/4\pi,$$

where $\Delta\phi_m$ is the measured phase difference. But since a phase measuring

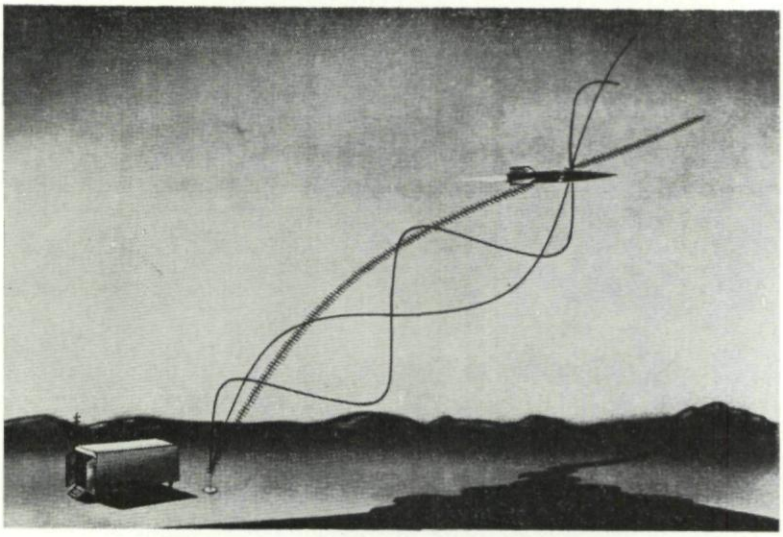


FIG. 3. *ITS-DME ranging techniques.*

device cannot distinguish among phase differences of $\Delta\phi$, $\Delta\phi + 2\pi$, $\Delta\phi + 4\pi$, a ranging signal of wavelength twice the expected maximum range must be used to avoid range ambiquities. For appreciable ranges, however, the precision of the range measurement is greatly reduced with the use of one long-wavelength ranging signal. The uncertainty introduced into an 800-mile range, for example, by a phase measuring device which can precisely resolve the phase difference to only one part in one-thousand (to approximately one-third degree), is about 4000 ft.

To obtain both the ambiguity-free range and the desired precision, the DME uses four ranging frequencies (.117 kc/s, 1.875 kc/s, 30.105 kc/s, and 480.234 kc/s). The very coarse frequency (.117 kc/s) allows the measurement of unambiguous ranges up to 795 miles, the coarse (1.875 kc/s) and the fine (30.015 kc/s) give progressively more precise measurements, and the very fine (480.234 kc/s) gives the desired precision to well within the specified two feet. Figure 3 is a simplified representation of this ranging technique.

These four ranging signals are combined and modulated on a UHF, carrier and transmitted to a target transponder where they are detected remodulated on a new UHF carrier, and retransmitted to the ground station. Upon return, they are compared in phase with the outputs of the signal generators and the ranging information is extracted. The ambiguities are resolved and a single ambiguity-free range measurement is calculated.

Once this range is obtained from each of at least three stations, spatial position and its time derivatives, velocity and acceleration, can be obtained.

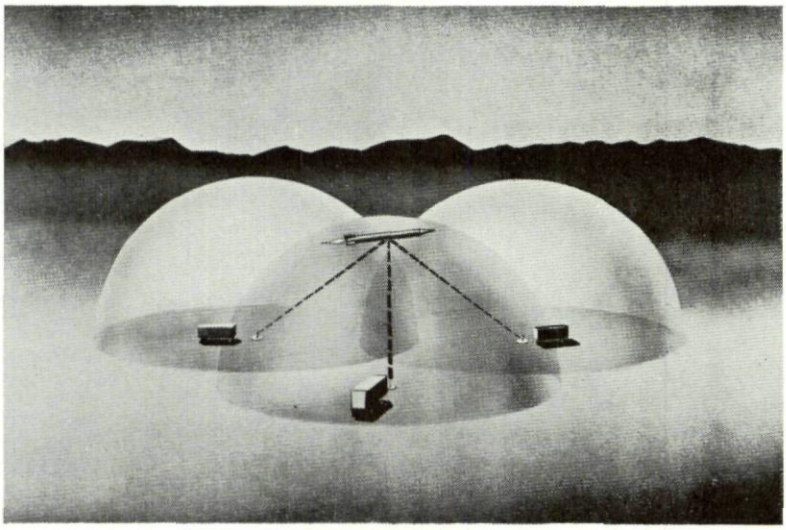


FIG. 4. *ITS-DME space position solution.*

A single range defines a sphere with the station at its center, and the intersection of three such spheres defines the spatial position of the target. (Actually the three spheres intersect at two points, but one is below the surface of the earth.) Figure 4 shows this spatial position solution.

This basic system can then be operated in one of at least two possible multiple-target modes:

(1) The targets, each equipped with a transponder of unique operating frequency, can be simultaneously interrogated by one of three sequentially operated ground stations.

(2) The targets, each carrying three transponders of unique frequencies, can be continuously interrogated by all three of the ground stations. The present DME system operates in the first mode and is capable of tracking up to three targets.

B. *Interferometry or Angle Measuring (AME)*

The AME is basically a radio interferometer. The phase difference between signals received at the ends of a known baseline is used to determine

the arrival angle of a CW signal from an airborne target. Neglecting atmospheric effects and assuming the station-to-target distance to be much greater than the baseline length, equal arrival angles at the receiving antenna can be assumed and the difference between the target-to-antenna paths is found to be

$$\Delta R = d \cos \theta,$$

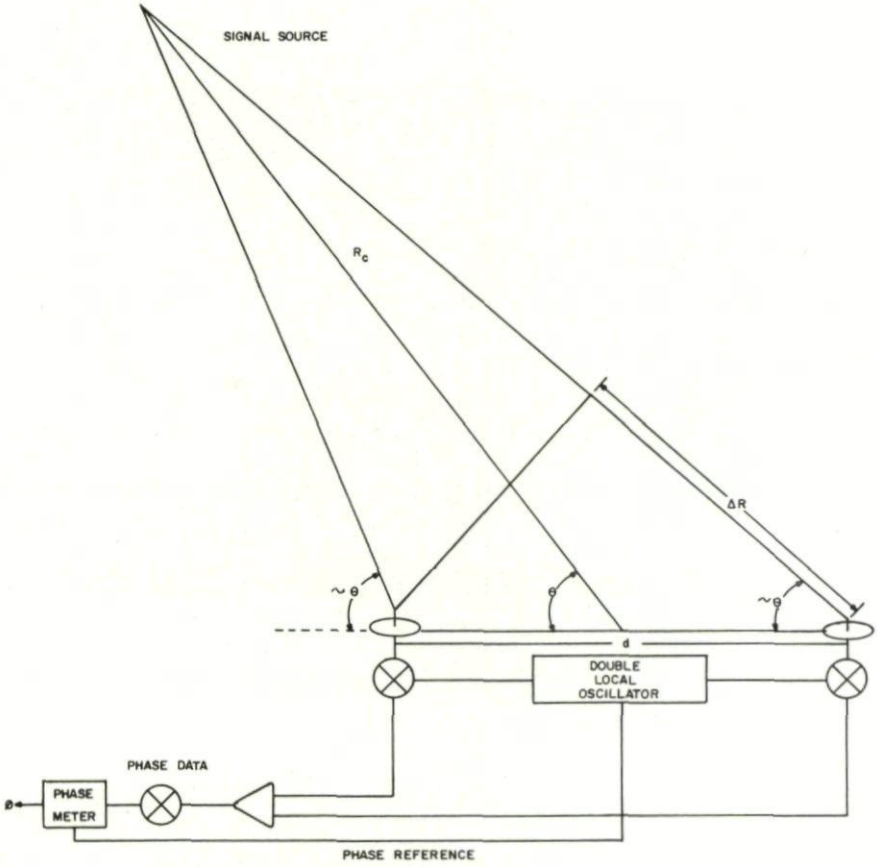


FIG. 5. *AME single antenna pair.*

where d is the baseline length and θ is the arrival angle. The measured phase difference between the antennas is, therefore,

$$\Delta\phi_m = 2\pi\Delta R/\lambda = 2\pi\, d \cos \theta/\lambda,$$

from which the arrival angle can be determined. (The arrival angle is never actually derived in the AME system, since the solution of the spatial position of a target requires the use of the direction cosine only.) Figure 5 illustrates the basic geometry of the AME.

To obtain a sufficiently precise, ambiguity-free angle measurement, the AME used four phase differences in a manner analogous to the resolution method of the DME. An arrangement of six antennas (on a common baseline) is used to obtain the phase differences corresponding to antenna separations of $.5\lambda$, 4λ, 16λ, and 128λ. The one-half wavelength separation gives the unambiguous angle measurement, the 4λ- and the 16λ-separations give progressively more precise measurements, and the 128λ-separation yields the ultimate precision of the system.

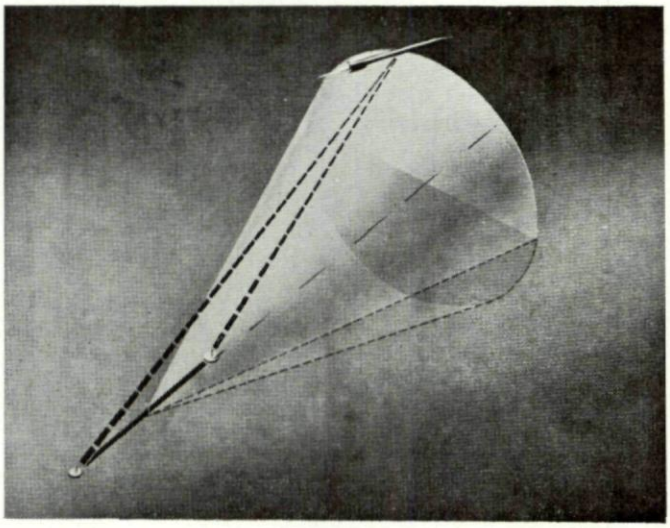


FIG. 6. *ITS-AME single antenna pair space position.*

Once the arrival angle is known, the target is fixed on the surface of a circular cone whose axis is the baseline and whose surface is defined by the arrival angle (Fig. 6). To exactly determine the direction to the target, each AME station used two perpendicularly crossed baselines of identically spaced antennas. Two intersecting cones are thereby defined, establishing as the target line, that line of intersection which extends above the earth. (See Fig. 7.)

Spatial position of the target is then fixed by the intersection of target lines from at least two remotely located AME stations. From these data velocity and acceleration can be determined. Figure 8 shows the system geometry for a given target position in relation to the existing AME sites

At least three modes of multiple-target tracking, each based upon the above basic system, are possible:

(1) Transmission by the target can be sequenced by a master ground station, allowing all ground stations to make measurements on the same frequency.

(2) Continuous transmission (at different frequencies) can be sequentially sampled, under master station control, by each of the ground stations.

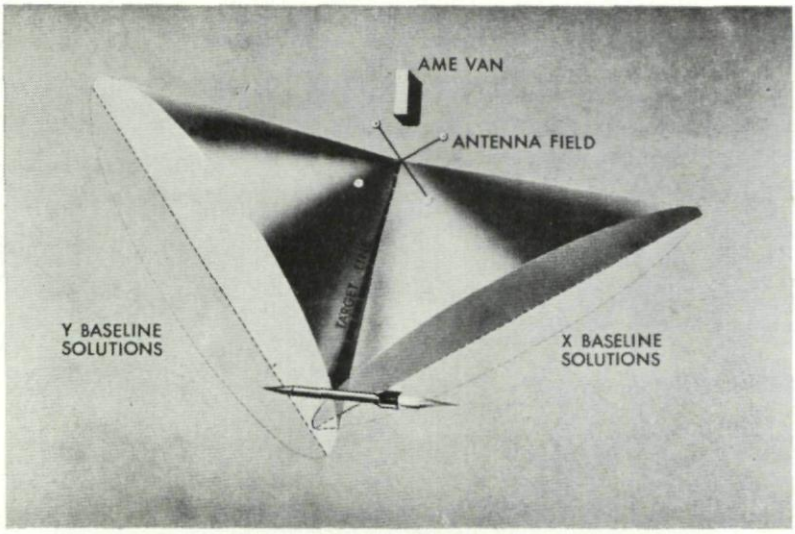


FIG. 7. *ITS-AME crossed baseline space position.*

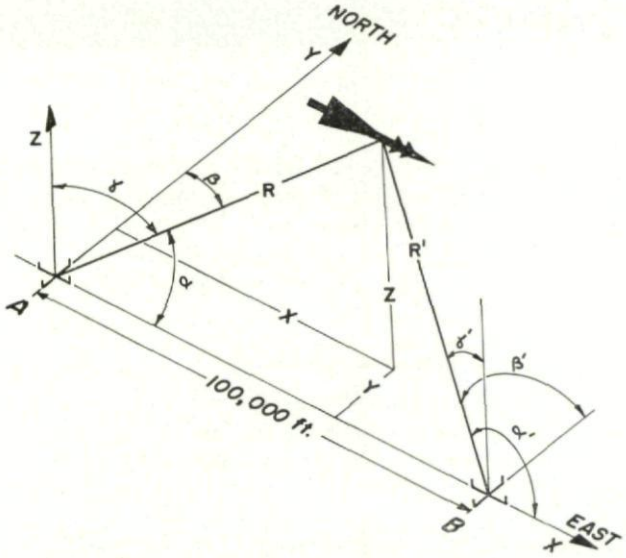


FIG. 8. *AME system geometry.*

(3) Continuous transmission (at different frequencies) can be continuously monitored by separate equipment at each of the ground stations. The existing AME system operates in Mode-3 and has a two-target capability.

C. *Doppler or Velocity Measuring (Doppler ME)*

The Doppler ME determines the rates of change of the ranges from one or more ground stations to a moving target by measuring the Doppler shift of a CW signal returned from the target. Neglecting atmospheric effects and assuming the general case of separated transmitter and receiver stations, the

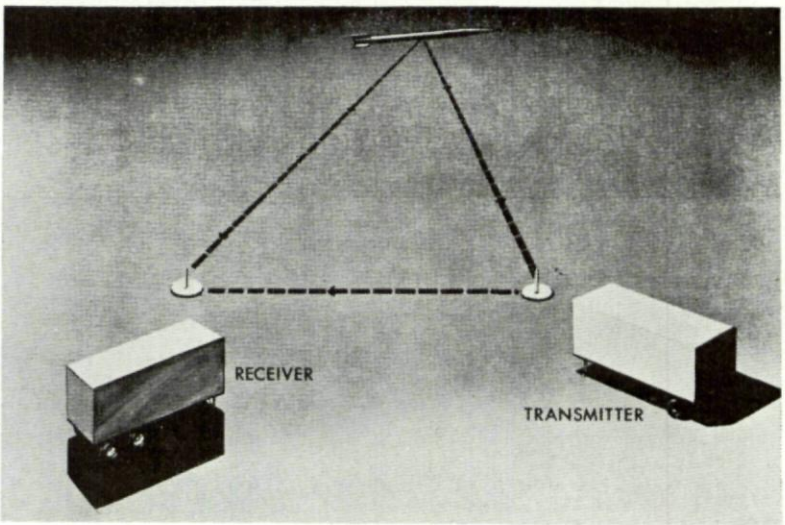


FIG. 9. *ITS doppler ME.*

Doppler shift of a signal transmitted to and received from a moving target is

$$f_D = f_0(\dot{R}_1 + \dot{R}_2)/c,$$

where f_0 is the frequency of the transmitted signal, c is the vacuum speed of light, and $\dot{R}_1$ and $\dot{R}_2$ are the rates of change of the transmitter-to-target range (R_1) and the target-to-receiver range (R_2), respectively. (In the degenerate case of a common transmitter-receiver station, the sum of $\dot{R}_1$ and $\dot{R}_2$ reduces to $2\dot{R}$, or twice the radial velocity of the target.)[1] The rate of change of the range sum ($\dot{R}_1 + \dot{R}_2$) is, therefore,

$$(\dot{R}_1 + \dot{R}_2) = f_{Dm}c/f_0,$$

where f_{Dm} is the measured Doppler shift obtained by comparing the signal received from the target with the signal received directly from the transmitter station.

At any given instant, it will be noted, the spatial position of the target lies somewhere on the surface of an ellipsoid whose foci are the transmitter and receiver stations and whose surface is defined by the instantaneous range sum $R_1 + R_2$. (See Fig. 9.) As the target moves, this ellipsoid can be thought of

as expanding and contracting so that the target continually rides its surface (where the rate of change of its major axis is equal to the range sum rate $\dot{R}_1 + \dot{R}_2$). The intersection (above the surface of the earth) of three such ellipsoids then defines the instantaneous spatial position of the target, and the instantaneous rates of change of the three ellipsoids define the target's velocity at that point. (Figure 10 shows the spatial position defined by the three intersecting ellipsoids.)

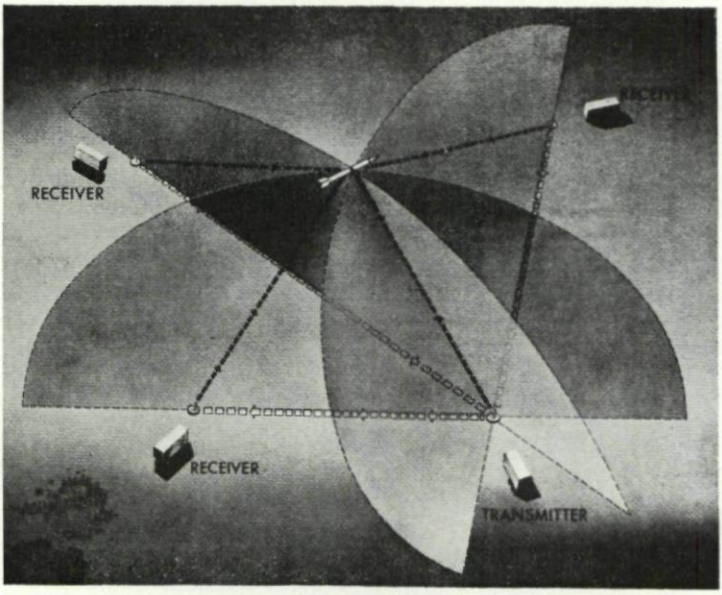


FIG. 10. *Three-station doppler solution.*

The instantaneous target velocity, however, cannot be determined in the absence of a known spatial position, since the instantaneous velocity which corresponds to three given range rates is also a function of the system geometry which exists at the time. And since a spatial position cannot be determined from three range rates alone, the system must be provided with known starting point (i.e. a known position such as the location of a missile launcher or a trajectory position as determined by another tracking system) to which the range rates can be referred. But when the starting point is once established, the range rates then provide sufficient information to allow subsequent target positions and corresponding velocities to be determined.

When the target moves from the start point to a subsequent trajectory point, each of the three range sums will be observed to change by the amount

$$\Delta R_s = N_i \lambda_0,$$

where N_i is the number of cycles of Doppler frequency recorded at the ith receiver during the interval of transition, and λ_0 is the wavelength of the reference signal. By applying these range sum changes to the known range

sums of the start point, the new spatial position of the target is established and the new velocity can be determined. Other spatial positions and velocities are, of course, obtained in the same manner.

The spatial positions and velocities of the proposed Doppler ME are expected to be mush more precise than those data which can be provided by either the DME or the AME; but due to the complicated reduction procedures which the Doppler ME data requires, these data will not be provided in real time. In addition, the system is being initially designed for a single-target tracking capability, and can therefore serve only as a supplement to multiple-target missions.

3. DESCRIPTION OF EQUIPMENT

3.1 *DME*

The ITS-DME[3] consists of at least three ground stations and as many as three target-borne transponders. Figures 11 and 12 are pictures of a typical

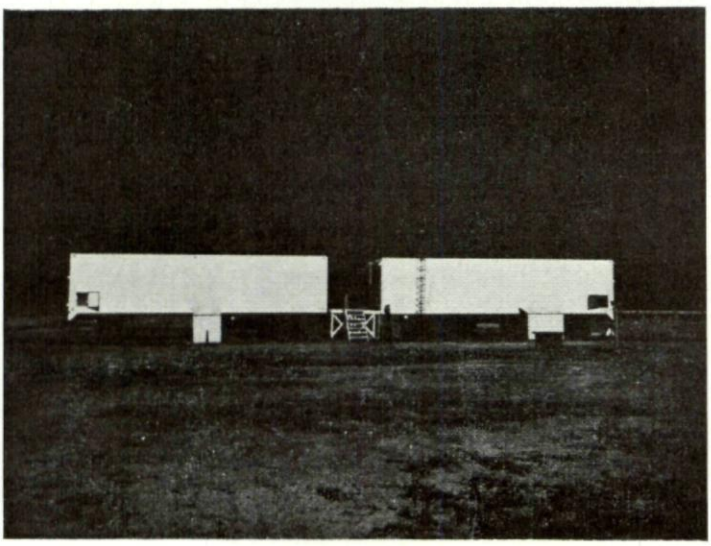


FIG. 11. *AME and DME ground stations.*

ground station and a typical transponder, respectively; and Fig. 13 is the block diagram of the DME system. The ground stations are identical and consist of four major subsystems: (1) the transmitter, (2) the receivers, (3) the phasemeters, and (4) the data handling equipment.

Figure 14 is a block diagram of the transmitter. The four ranging frequencies (.117 kc/s, 1.875 kc/s, 30.015 kc/s, and 480.234 kc/s) are generated in the frequency synthesizer shown in the above block diagram. The frequency of highest resolution, 480.234 kc/s, was selected to provide instrumental resolution of $\pm$ one foot, based upon the ability of a phasemeter to

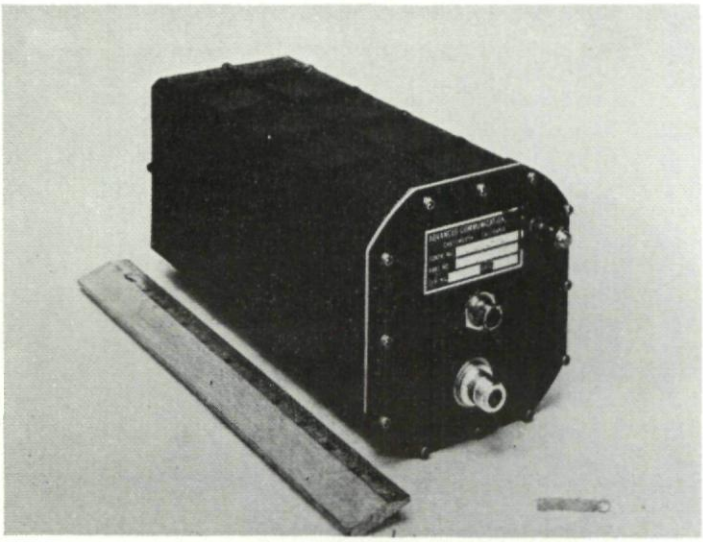


FIG. 12. *ITS-DME transponder.*

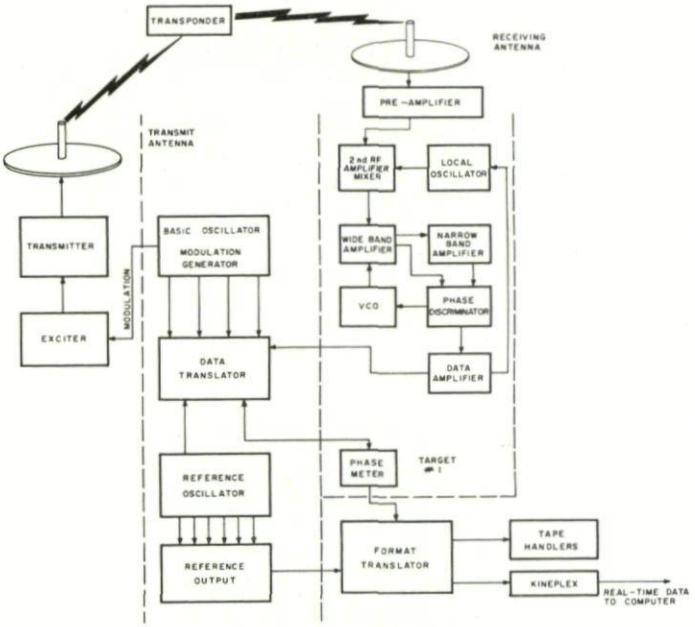


FIG. 13. *ITS-DME block diagram.*

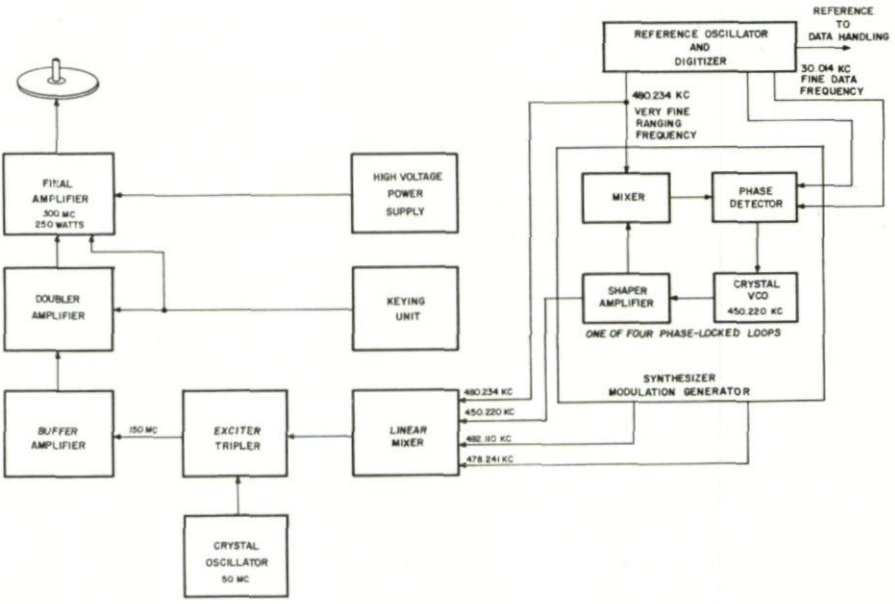


FIG. 14. *DME transmitter.*

read to $\pm 1/3$ degree. (Table 1 gives the instrumental resolution and maximum range of each of the ranging frequencies). Each ranging frequency is related to the next higher and/or lower frequency by a factor of 16, a factor chosen for data handling convenience. To avoid a modulation bandwidth of almost 480 kc/s, the ranging frequencies are translated into the relatively narrow bandwidth of about 30 kc/s, centered about 465 kc/s. The translation process is accomplished with phase-locked loops to maintain phase coherence, and consists of linearly mixing the four frequencies. These new frequencies are then used to modulate the transmitter with a deviation ratio of 0.3.

TABLE 1

RELATIONSHIP OF PRECISION AND RANGE TO FREQUENCY OF DATA SIGNALS

Frequency f	Wavelength	Maximum unambiguous range (one-way distance)	Maximum precision	Transmitter modulation	
				Mult. factor	Frequency
117.5 c/s	1585.00 miles	795.00 miles	4,096 ft	4079/4096	478.241 kc/s
1876.0 c/s	99.2 miles	49.6 miles	256 ft	252/256	482.109 kc/s
30.014 kc/s	6.21 miles	3.1 miles	16 ft	15/16	450.219 kc/s
480.234 kc/s	0.39 miles	0.19 miles	1 ft	1	480.234 kc/s

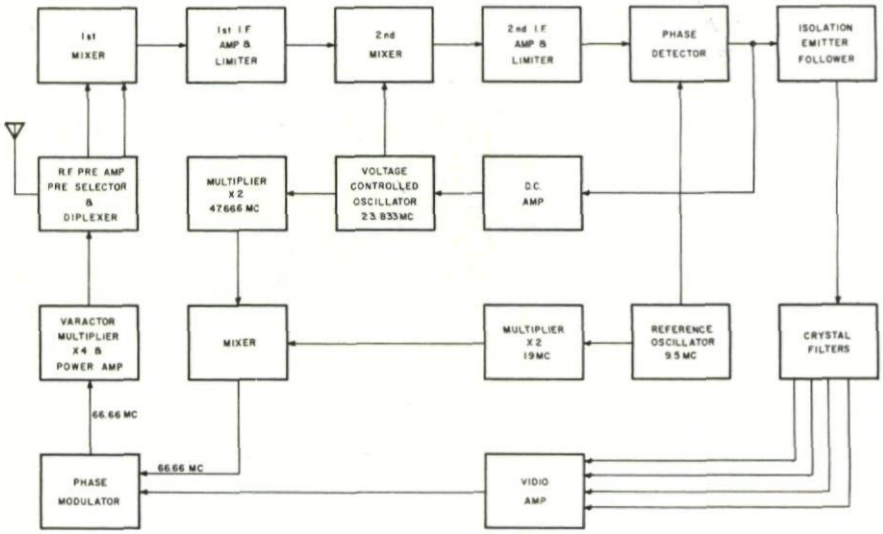


FIG. 15. *ITS-DME transponder block diagram.*

The transmitter is crystal controlled on a 300 Mc/s carrier, and nominally radiates 250 W from a single quarter-wave "stub" antenna. To prevent the simultaneous interrogation of the three transponders, the three DME transmitters are sequentially keyed, where the time sequence is controlled by central range timing. Each transmitter is left on sufficiently long for the signal to traverse the maximum system design range of 400 miles and to return.

Figure 15 is the block diagram of a DME transponder. In the transponders (all of which receive the 300 Mc/s carrier), the ranging frequencies are extracted, the carrier is fractionally multiplied to one of three lower frequencies, and the ranging frequencies are used to modulate the new carrier, which is then amplified and retransmitted to the ground stations. (Table 2 shows the relationship between the received and retransmitted carriers.) Phase coherency of the carrier and the four ranging frequencies is maintained throughout each of the transponders.

TABLE 2

CARRIER RELATIONSHIP IN DME TRANSPONDERS

Transponder	Input frequency	Multiplication factor	Output frequency
Number 1	300 Mc/s	8/9	266.666 Mc/s
Number 2	300 Mc/s	9/10	270.000 Mc/s
Number 3	300 Mc/s	12/13	276.923 Mc/s

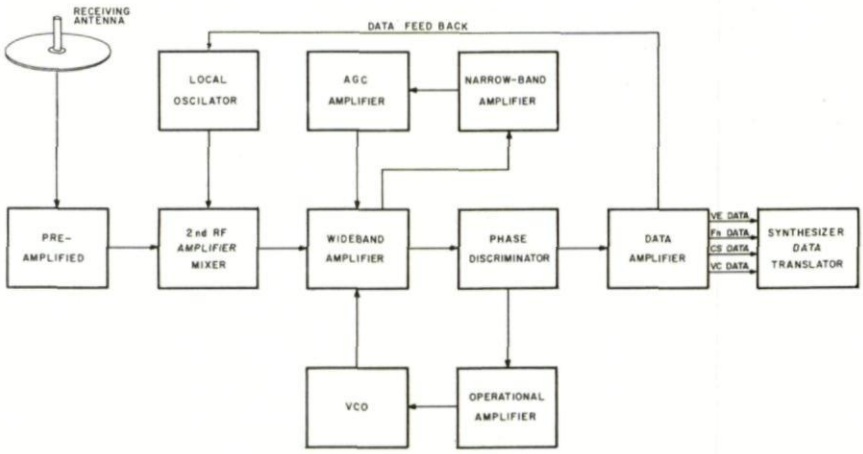


FIG. 16. *ITS-DME receiver.*

At the ground station a broadband preamplifier accepts the three new carriers from another quarter-wave stub. The ranging frequencies are detected in a conventional phase-locked loop demodulator, and are then separated in crystal filters and amplified for processing in the phasemeters Figure 16 is a block diagram of the DME receiver.

Figure 17 is a block diagram of a digital servo phasemeter, four of which are used for each receiver (i.e. a total of twelve per ground station). The four

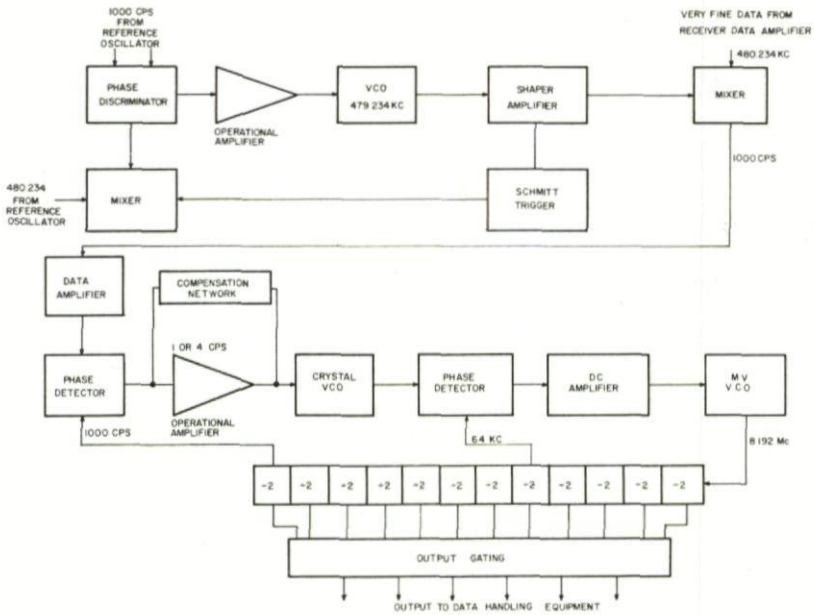


FIG. 17. *Electronic phasemeter simplified block diagram.*

ranging frequencies are translated to 1 kc/s, a process in which phase-locked loops are again used; reference ranging frequencies from the synthesizer and a 1 kc/s reference frequency derived from an 8.192 Mc/s source are also used. The slave counter is gated on to count the 8.192 Mc/s from the VCO, its on-period being controlled by the phase difference between the return and reference signal phases. The range is thus generated in digital form, where the outputs of each of the four phasemeters is a 13 bit binary word. Digitized range information is then processed in the data handling equipment to make it suitable for recording on magnetic tape at the ground station, or for transmission to a digital computer for the real time display of spatial position.

3.2 *AME*

Figure 18 is a block diagram of the AME. The AME consists of two remotely located ground stations, each of which is capable of receiving any two

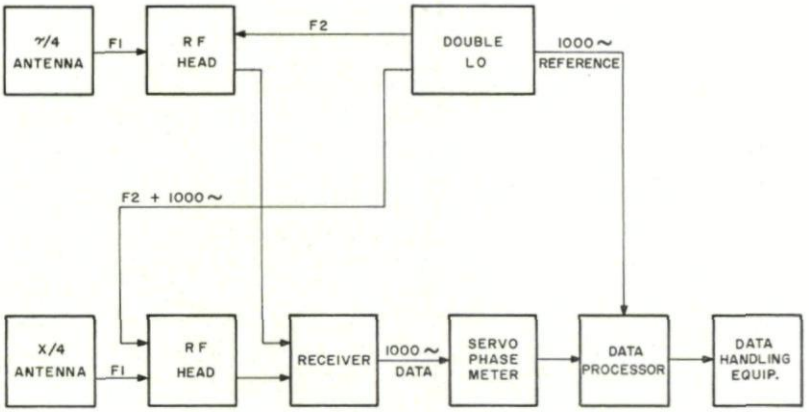


FIG. 18. *AME block diagram.*

frequencies between 215 and 277 Mc/s. These signals are typically 500 kc/s bandwidth telemetry carriers, but any one of the three DME transponder carriers can also be received. To avoid mutual coupling between antennas and to reduce the number of antennas required on each baseline, all but the highest resolution pair are located nonsymmetrically, as shown in Fig. 19. By properly subtracting phase differences (in the processor), data from the equivalent of four antenna pairs are obtained. (Table 3 gives the resolution and maximum unambiguous angle for each of the antenna pairs.) It is recognized that a slight error is introduced into the data by this non-symmetry, but since the system accuracy is established by the 128λ antenna pair and not by the non-symmetrically arranged antennas, the error has no real effect on the final result.

The signal from each antenna is square law detected; and the signals which correspond to an antenna pair are phase compared in a phasemeter identical to those used in the DME. An unambiguous direction cosine for

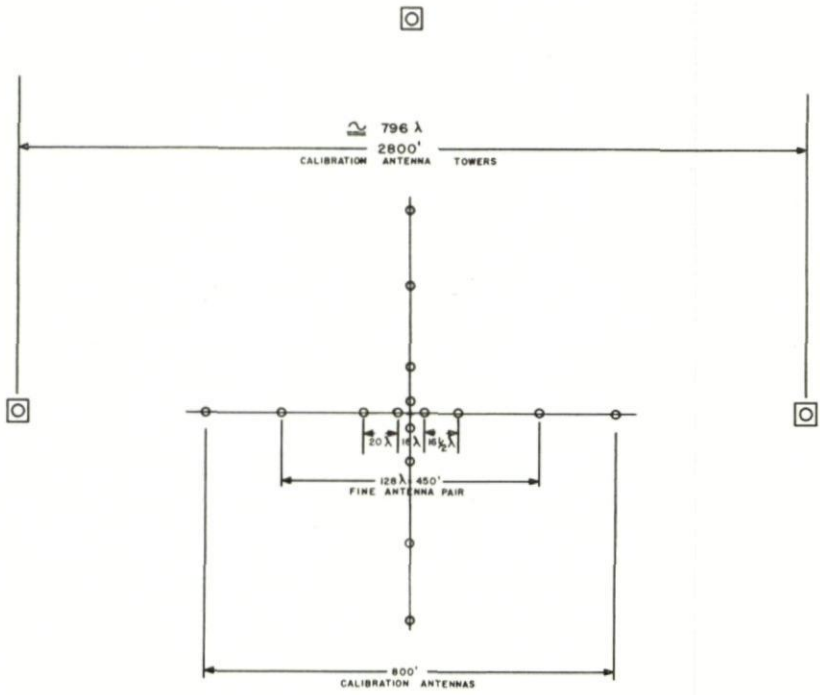


FIG. 19. *ITS-AME antenna field.*

TABLE 3

Baseline length at 277 Mc/s 454.356 ft	Angle resolution	Unambiguous angle
128 λ	6′ to 3″	0.70325°
16 λ	0.022°	5.625°
4 λ	0.087°	22.5°
05 λ	0.74°	0–180° (– 90° to 90°)

[a] 00° 06′ 00″ to 00° 00′ 03″ depending on value of cosine

each of the two targets is then computed, in digital form in the AME processor, for each of the baselines. The digital data handling equipment is very similar to that of the DME, and will not be discussed.

3.3 *Doppler ME*

It was mentioned earlier that the Doppler ME is currently being developed. One master transmitter and six slave stations are being constructed;

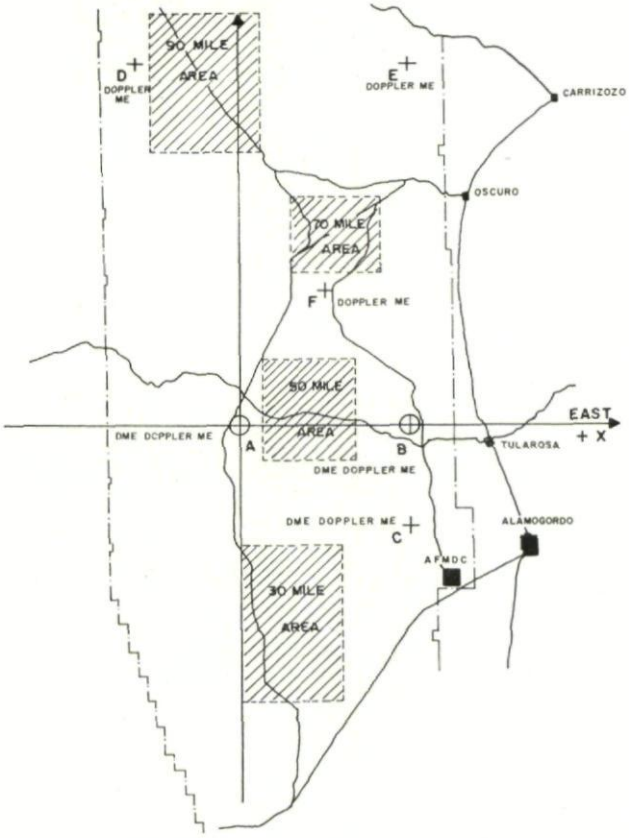


FIG. 20. *ITS Doppler ME configuration.*

three of the slave stations will be integrated into the three existing DME stations and three are being designed to operate independently. Figure 20 shows the planned operational configuration, but, it should be noted, additional stations may be added to the system as required. Figure 21 is a block diagram of the DME-Doppler ME station, and Fig. 22 is a block diagram of an independent Doppler ME station.

The Doppler ME uses, as a reference, a master transmitter so located that there is rf line of sight between it and each of the six slave stations. (See

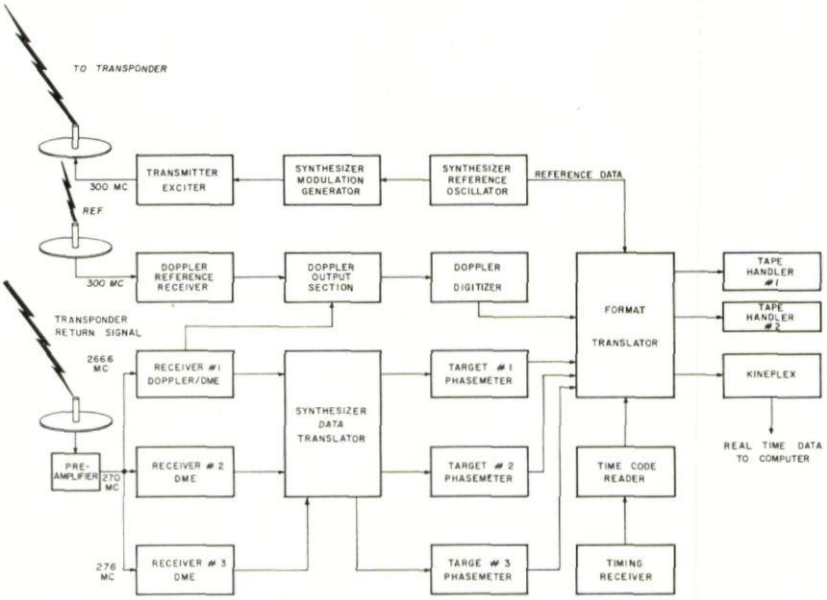


FIG. 21. *ITS Doppler DME.*

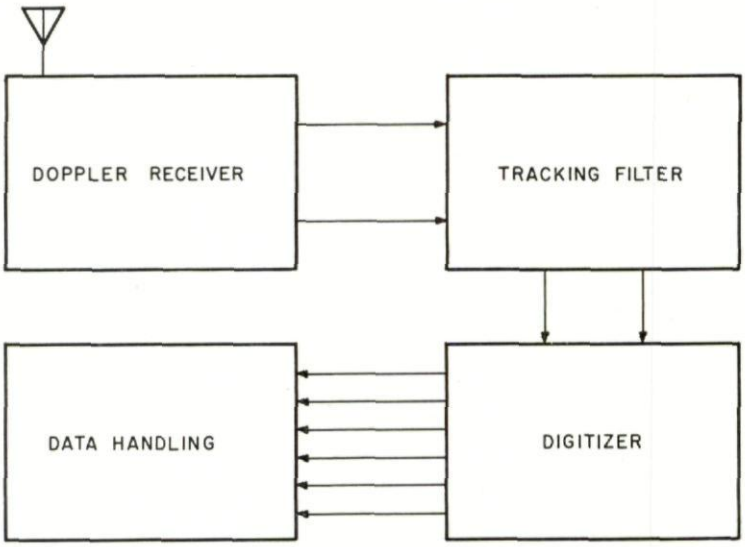


FIG. 22. *ITS Doppler ME ground station.*

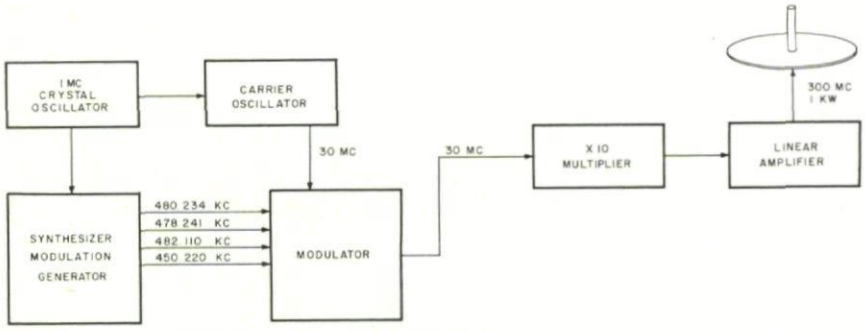


FIG. 23. *ITS DME-Doppler master transmitter*

Fig. 23 for the block diagram of the transmitter). The master station transmits a 300 Mc/s signal which may or may not be sequenced for the simultaneous operation of the DME. This signal is received by each of the six slave stations which use it as a reference, and by the transponder (Fig. 24) where it is fractionally multiplied by a factor of 8/9 to give 266.6 Mc/s. (A DME transponder may also be used for this.) The modified carrier is retransmitted to the six stations, where the Doppler frequency is detected by phase comparing the reference signal with the retransmitted signal. This information is then digitized and recorded on magnetic tape. The data handling equipment, however, does not have the capacity for processing the Doppler ME data for real time transmission, but this capability is planned for later implementation.

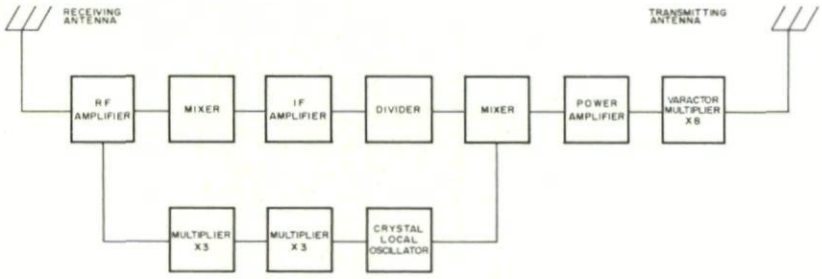


FIG. 24. *ITS Doppler ME transponder.*

4. PERFORMANCE ANALYSIS

4.1. *Demonstration System Performance*

Three tables, Tables 4, 5, and 6, delineate the instrumental characteristics of the ITS-DME, -AME, and -Doppler ME. It is seen that between conceptual design and implementation, a number of requirements had to be compromised.

Calibration of the system also proved to be more difficult than expected. In the DME, anomalous and variable range errors were detected in the

TABLE 4

DME CHARACTERISTICS

Equipment	Design	Actual	Planned
Transmitter			
Ranging Frequencies	478.241 kc/s	Same	Same
	482.110 kc/s	Same	Same
	450.219 kc/s	Same	Same
	480.234 kc/s	Same	Same
Ranging Frequency Stability	±2 c/s	Same	Same
Phase Stability	0.3°	Same	Same
Transmitter Frequency	300 Mc/s	Same	Same
Transmitter Stability	One part in 10^7/hr	Same	One part in 10^{10}/hr
Transmitter Output Power	200 W	Same	1000 W
Transmitter Antenna Gain	unity (2/4 stub)	Same	Same
Receiver			
Preamp Noise Figure	5 db	Same	Same
Preamp Bandwidth	20 Mc/s	Same	Same
Receiver Sensitivity	− 120 dbm (carrier)	Same	Same
Receiver Bandwidth	4 Mc/s	Same	Same
Receiver Frequencies	266.666 Mc/s ±12 kc/s	Same	Same
	270.000 Mc/s ±12 kc/s	Same	Same
	276.923 Mc/s ±12 kc/s	Same	Same
Phasemeter			
Resolution	1/27°	Same	Same
Stability	1/3°	Same	Same
Data Handler			
Output Word	IBM Format	Same	Same
Input Capacity	120 Words	Same	Same
Output Word	36 Bits	Same	Same
Transponder			
Transmitter Frequency	300 Mc/s	Same	Same
Receiver Frequencies	266.666 Mc/s	Same	Same
	270.000 Mc/s	Same	Same
	276.923 Mc/s	Same	Same
DME Accuracy characteristics			
Slant Range Measurement	±2 ft	Same	Same
Ground Station Antenna Survey	First Order	Same	Same

calibration process. (Calibration was done by comparing the measured range from the station to a target located on a mountain with the same range derived from surveys.) In addition to the above bias errors, there were also random errors that were attributed to variations in refractive index along the propagation path. This type of error is known to be particularly high at low elevation angles, as in the calibration configuration mentioned. The calibration problem for the DME is not yet completely solved.

In the AME case, calibration antennas were placed along the crossed baselines near the operating antennas. Variations in calibration measurements

TABLE 5

AME CHARACTERISTICS

Very fine antenna spacing	±0.01 ft
Parallelism with coordinate axis	±3 sec arc
Distance accuracy between sites	±1 part/100,000
Parallelism between sites	±3 sec arc
Accuracy of cosine measurement (before refraction correction)	±50 parts/1,000,000
Accuracy of cosine measurement (with corrections)	±15 parts/1,000,000
Receiver Noise Figure	5 db
Receiver Bandwidth (max)	65 Mc/s
Receiver Sensitivity	— 133 dbm

were detected, however, and were attributed to the variable phase pattern in the near field of the calibration antennas. To solve the problem they were moved to 400 wavelengths from the operating antenna field and elevated to 100 foot towers. Another problem encountered was that the equipment vans were located close enough to the antenna field to cause receiver interference from multipath reflections. The vans were therefore moved 1000 feet away from the field, and apparently the multipath problem has been eliminated. But in so doing, one of the system design characteristics had to be eliminated also, viz. the common DME-AME antenna field which would have permitted a biconical-spherical position solution, since it would be impractical to send 250 W of power through the 1000 feet of coaxial cable from the van to the antenna field. This problem may be resolved in the future, though, by locating the equipment underground near the antenna field.

TABLE 6

DOPPLER DME CHARACTERISTICS

Transmitter	
Frequency	300.000 Mc/s
Frequency Stability	1 part in 10^{10} per day
Phase Stability	1° in 80 hours
Data Receiver	
Frequencies	266.666 Mc/s ±12 kc/s Target #1
	270.000 Mc/s ±12 kc/s Target #2
	276.923 Mc/s ±12 kc/s Target #3
Reference Receiver	300 Mc/s ±12 kc/s

There are, of course, no performance data for the Doppler ME which is still under development.

4.2 *Geometrical Dilution of Precision (GDOP)*

Geometrical dilution of precision, as the name implies, is the degradation by the system geometry of the determination of a spatial position from imprecise measurements.

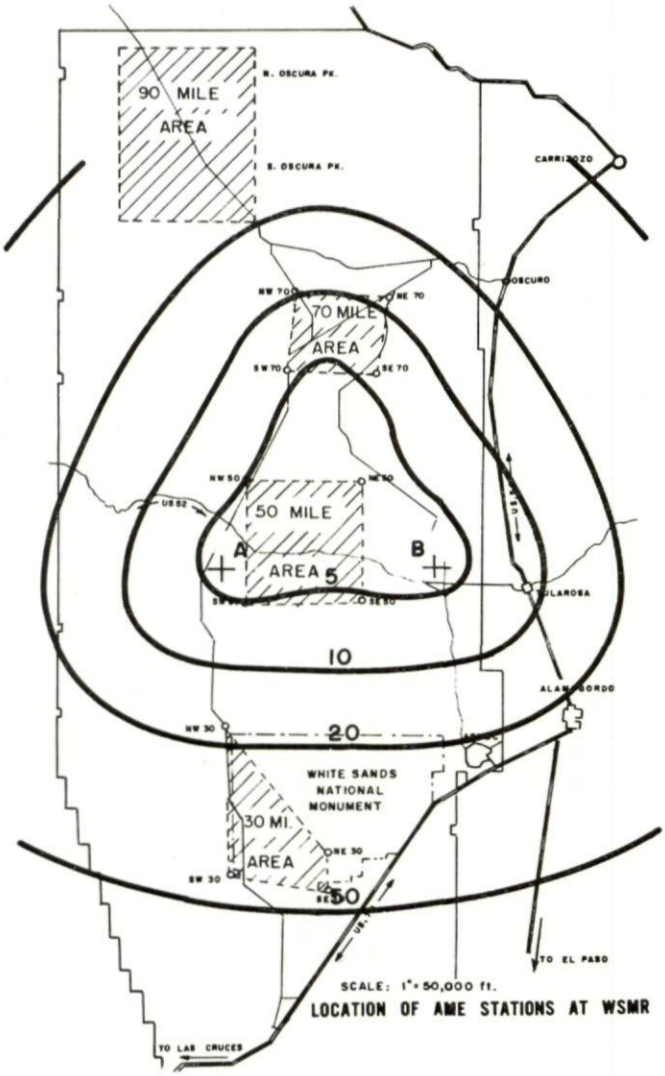


FIG. 25. *DME Error contour* $Z = 25{,}000$.
Total position error

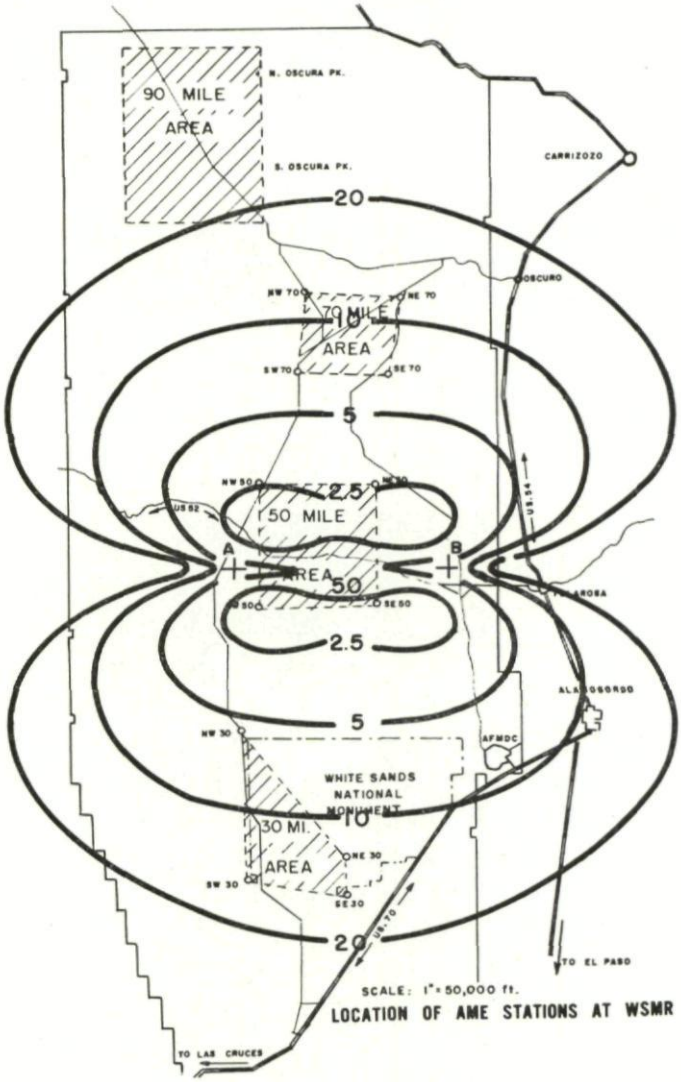


FIG. 26. *AME error contour* $Z = 20{,}000$.
Total error for least squares solution

GDOP was calculated for the DME system,[5] and equal-error contour curves were drawn for the three station position estimate. Figure 25 is one of these error contour curves. For example, in an area above the 70 mile area, the total position error due to DGOP for a target at an altitude of 25,000 feet is between 5 and 10 feet. The particular ground station configuration chosen was an equilateral triangle, and the error of observation in distance measurement at each station was assumed to be 2 feet (a design specification for the

DME). For this configuration, the position of least GDOP is along a line extending vertically from the center of the equilateral triangle formed by the stations.

GDOP was also calculated for the AME system,[6] and equal-error contour curves were drawn for a two-station AME position estimate. Figure 26 is one of these calculated AME error contour curves. For example, at a range of about 25 miles north from the antenna fields, the total space position error due to GDOP alone is about 20 feet, where the error of observation of direction cosine at each station was assumed to be 15 ppm (the design specification for the AME). The position of minimum GDOP is variable, but generally exists in the volume above the plane area surrounding and connecting the two AME stations; at 100,000 feet and above, this volume of minimum GDOP reduces to the line that is the perpendicular bisector of the line connecting the two stations.

An error analysis of the Doppler ME has not been undertaken, but such an analysis will be similar to the one already performed on the DOVAP system.[14]

5. CONCLUSIONS

Operational feasibility of the Integrated Trajectory System has been demonstrated. Although the equipment field performance has not been up to expectations, it is believed that the principal system problems have now been well defined, and that the solutions to these problems will be forthcoming in the foreseeable future. It is also reasonable to expect that ITS will be expanded and improved to provide the high performance required at WSMR.

6. DISCUSSION

W. Djinis: What is the capability of range resolution of the DME technique which you have used?

G. F. Bigelow: The range resolution is about 2 feet. This limit is set partly by the digital data handling equipment and partly by atmospheric effects at low angles of elevation.

W. Djinis: How does this vary with the fractional multiplication of the output when more than one target is being tracked?

G. F. Bigelow: The precision is reduced, probably by a factor of five.

W. Djinis: Is the range jitter of the transponder insignificant compared with the 2 foot range resolution of the DME?

G. F. Bigelow: Yes the range jitter is insignificant with the present transponder. Incidently the transponder jitter is low even under the conditions of missile environment, which is important for this work.

REFERENCES

1. Perkins, G. O., *Proposal for an Integrated Trajectory System.* Range Instrumentation Division (1958).
2. Miller, V. L., *Electronic Trajectory Instrumentation State of the Art.* Range Instrumentation Development Division, 1960.

3. *Preliminary Handbook for Integrated Trajectory System Distance Measuring Equipment*, Vol. 1, CUBIC Corporation.
4. *Preliminary Handbook for Integrated Trajectory System Distance Measuring Equipment*, Vol. II, CUBIC Corporation.
5. WEISS, J. E. *Error Contour Curves for a DME/DME Configuration*. Range Instrumentation Development Division, 1961.
6. HOLDBERG, D. E., and MILNARICH, P., Jr., *Standard Error Analysis of Angle Measuring Equipment as Part of the Integrated Trajectory System at White Sands Missile Range*, Vol. I. Range Instrumentation Development Division, 1960.
7. HOLDBERG, D. E. and MILNARICH, P., Jr., *Standard Error Analysis of Angle Measuring Equipment as Part of the Integrated Trajectory System at White Sands Missile Range*, Vol. II. Range Instrumentation Development Division, 1960.
8. MILNARICH, P., Jr., *A method for Predicting Apparent Elevation Angle for a Tropospheric Target from Specified Values of Geometrical Elevation Angle and Height*. Instrumentation Department, USA ERDA-67, 1963.
9. PAUL, R. H., *Some Effects of the Atmosphere and Geometry on the Radio Interferometer Angle Measurement*. Instrumentation Department, SELWS-E-63-12, 1963.
10. HOLDBERG, D. E., *DME Data Reduction*. Instrumentation Department, USA ERDA-1, 1963.
11. VOSS, R. A. and PAUL, R. H., *An Approach to the Determination of the Effective Lengths of Reference Signal Paths in the DME-Doppler System*. Instrumentation Department, USA ERDA-77, 1963.
12. PAUL, R. H. and MILNARICH, P., Jr., *Simultaneous Refraction Error Removal and Atmospheric Refractivity Extraction*. Range Instrumentation Development Division Interim Report Nr. 1, 1962.
13. WHITE, J. E., *The Least Squares Procedure Applied to the Problem of Estimating Missile Position and Atmospheric Refractive Index from DOVAP Range Data*. Instrumentation Department, USA ERDA-43, 1963.
14. Final Report, *Doppler Velocity and Position Study*. Aeronutronic Systems, Inc. 1959

CHAPTER 18

METHODS FOR OBTAINING VELOCITY AND RANGE INFORMATION FROM CW RADARS

MAHLON EASTERLING
Jet Propulsion Laboratory, Pasadena, California, U.S.A.

Continuous wave radars are particularly suited for obtaining velocity and range information on single targets at extreme ranges whether they be spacecraft in orbit about the earth or en route to the Moon or another planet, or natural celestial bodies such as the Moon or inner planets. The carrier itself provides the means for determining the velocity of the target because of the Doppler shift of the returned signal. Ways of measuring this Doppler shift for both coherent and noncoherent returned signals are discussed. The determination of range requires that the carrier be modulated. Methods of deriving modulating waveforms from pseudo-random sequences which are suitable for achieving high resolution and accuracy and resolving ambiguity are given, as well as ways of extracting the range from a returned signal. Techniques are described for both the coherent and noncoherent carriers. The several techniques described have all been embodied in actual radar systems and used to track appropriate targets. Experimental results are given to show the performance which can be attained.

1. INTRODUCTION

Continuous wave (CW) radars are particularly suited to tracking spacecraft because the same carrier can also be used for communications. The presence of a carrier allows narrow band detection methods to be applied to both functions. A factor that has contributed to the rapid development of these radars is that their operation is inherently the same with or without a transponder. Therefore, much of the development work has been done using passive reflectors, for example the Moon and inner planets.

The most useful kinds of tracking information which can be obtained by radar from very distant targets are radial velocity and range. The carrier in a CW radar provides a direct means for determining the radial velocity by measuring the Doppler shift. To obtain range information, it is necessary to modulate the carrier and use the returned modulation to determine the round trip propagation time for the signal. This *time of flight* is related to the range by the propagation velocity.

The CW radars to be discussed may be divided into two classes according to whether or not the returned signal is coherent, or, more to the point, whether it can be treated as coherent by the receiver. In the discussion which follows the velocity measuring methods are treated first, both the coherent

and noncoherent cases. Results of actual tracking operations are given as each method is discussed. The treatment of methods for measuring range begins with a development of the theory of the modulating signal derived from pseudo-random sequences. This is followed by methods applicable to the coherent radar and to the noncoherent radar. Results are given for each method as it is presented. Since many of the techniques presented here were developed at the Jet Propulsion Laboratory (JPL) for the National Aeronautical and Space Administration (NASA)/JPL Deep Space Instrumentation Facility (DSIF), the descriptions of the techniques apply to an L-band or S-band radar, and the results are also for these frequencies. However, by a suitable adjustment of parameters, the techniques are applicable at other frequencies.

2. VELOCITY MEASUREMENT WITH A COHERENT CW RADAR

A Doppler measuring coherent CW radar is shown in Fig. 1. in simplified form. Note that, in this figure and in others to follow, many absolutely vital parts of the radar are omitted. There will be no discussion of antennas, feeds,

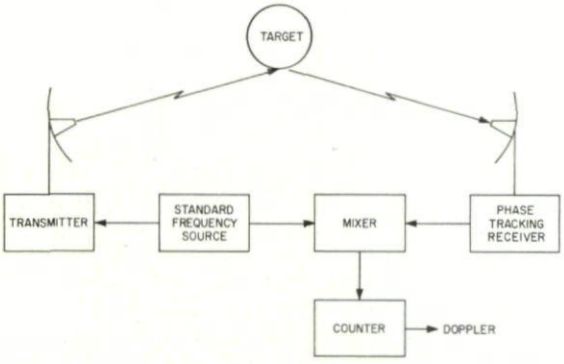


FIG. 1. *Doppler measurement by Coherent CW radar.*

masers, transmitters, and other parts which are both necessary and of great interest. Suffice it to say that appropriate devices of this sort were used in the examples to follow and were necessary for the overall operation of the system.

In Fig. 1, it is assumed that the target is passive. A CW signal derived from a frequency standard is transmitted to the target and reflected to the receiver. The receiver tracks the phase of the received signal and produces an output which is a filtered version of the input. This output, which is shifted in frequency by the Doppler effect, is mixed with the transmitted frequency and the difference frequency counted. Since the receiver tracks the phase of the returned signal there is no frequency error as such; errors due to noise appear as jitter in the measurements, but there is no bias. This permits statistical methods to be applied to smooth the measurements.

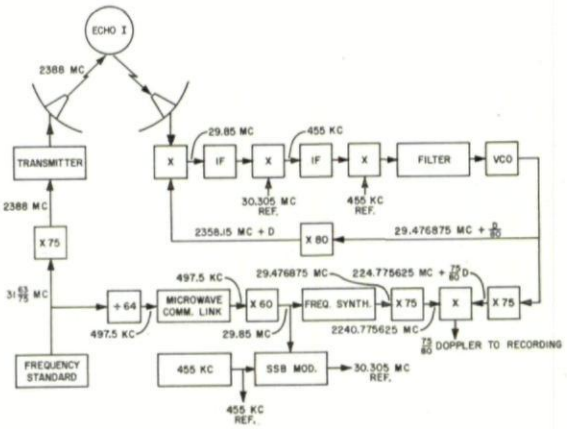


FIG. 2. *Velocity tracking system used for Echo* 1.

An early application of this technique was in the tracking of the Echo I balloon satellite[1,2] and in tracking experiments on the Courier satellite.[3] In regard to the Courier satellite, a full tracking and orbit determination experiment was performed, but the target was treated as a passive device; no use was made of the on-board electronics. A block diagram which shows the form of the receiver and the way in which the mixing is done is given in Fig. 2.

In this system, the two antennas were separated by some 7 mi and two ranges of mountains to provide isolation. A microwave communications link was used to provide a coherent reference for detecting the Doppler shift on the received signal. The receiver was a double superheterodyne phase-locked loop. The extra heterodyning signals are derived from the coherent reference and a free running 455 kc/s oscillator, but the free running oscillator is used in such a way that its frequency does not enter into the output. The actual Doppler as detected and counted was 75/80 of the Doppler on the returned signal. The somewhat unusual frequencies used in the system were selected for ease of mechanization and to combat leakage problems.

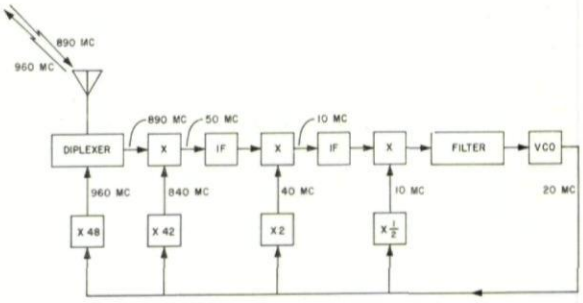


FIG. 3. *Coherent transponder.*

The system shown in Fig. 2 is inherently capable of detecting a change in propagating path length of one wavelength of the carrier (times 80/75) or about 13.5 cm. If the Doppler frequency is counted for 1 s the velocity resolution is 6.75 cm/s. If the frequency is counted for a longer time the resolution theoretically increases. Actually, the useful resolution is limited by the stability of the system and the accuracy and stability of the frequency standard. In tracking the Courier satellite, the Doppler was counted for 1 s every 2 s and an attempt was made to determine the accuracy of the tracking data by comparing them with a best fit orbit. The rms velocity

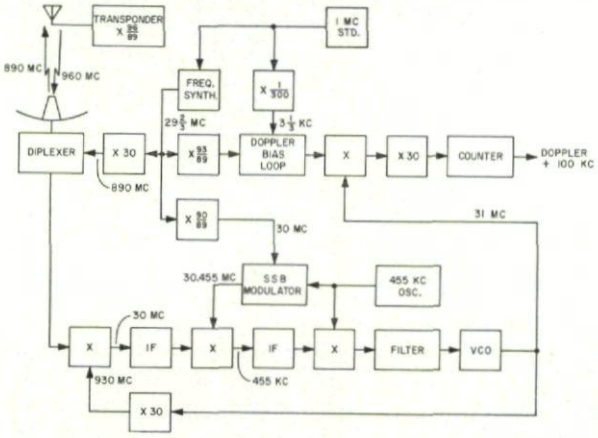


FIG. 4. *Doppler measurement with a coherent transponder.*

residual compared to the orbit was 9.4 cm/s but other evidence indicates that a part of this discrepancy was due to the orbit determination program. Any mathematical model used for computing an orbit necessarily contains many simplifying assumptions and a number of constants pertaining to the geometry of the situation. The velocity residual contains these effects as well as the effects of errors in tracking.[3]

The overall Doppler measuring system is shown in Fig. 4. The transmitter and receiver are very much the same as those in Fig. 2 except that different frequencies are used. The Doppler measuring system is different from that previously shown in two ways. One difference is that the signal is first detected at the VCO frequency and then multiplied to give the Doppler instead of being multiplied before detection. The result is the same. The other difference is that the Doppler frequency is deliberately biased by introducing an offset into the reference to facilitate counting, since the counted frequency never becomes zero. There is no effect on the accuracy of the system.[4]

When the *Mariner II* spacecraft to Venus was being tracked, the usual frequency standard at the tracking stations was a high quality crystal oscillator. This standard proved to be the limiting factor in the accuracy of the

Doppler measurement. The Doppler was counted for 50 s each 60 s. When the propagation time to the spacecraft and back was 20 s, the rms error in velocity compared to a best fit orbit was 7.5 cm/s. One tracking station, Goldstone, was fitted with a rubidium vapor frequency standard. When the propagation time to the spacecraft and back was 58 s, the rms error in velocity compared to a best fit orbit was only 0.3 cm/s. This extreme accuracy can now be obtained in routine operations.

The technique used to measure the velocity of a passive target can be extended to an active target, i.e. a target which includes a transponder. Ideally, the transponder would be a reflector with power gain. Practically, it is necessary to receive on one frequency and transmit on another. This can be accomplished by using a phase tracking receiver in the transponder and deriving the transmitter signal from the output of the receiver. If, in addition, all heterodyning frequencies are also derived from the output of the receiver, the transmitted signal is coherent with the received signal. The frequency of the transmitted signal is a constant times the frequency of the received signal. This apparent increase in complexity can be turned to advantage in the radar itself. Since the transmitter and receiver are at different frequencies, they can be diplexed on the same antenna. Since the antennas are large and expensive, this is a distinct advantage. The disadvantage is that the Doppler equation is more complex, but since the tracking data are usually reduced by large computers, this is not important.

The *Ranger* and *Mariner* spacecraft have used a coherent transponder of the form shown in Fig. 3. The output frequency is exactly 96/89 times the input frequency.

3. VELOCITY MEASUREMENT WITH A NONCOHERENT CW RADAR

The noncoherent CW radar has been developed to overcome certain limitations of the coherent CW radar. These limitations are of two kinds; there is a minimum signal level which can be tracked, and the coherent receiver is inherently unsuited to the reception of certain kinds of noncoherent signals. The receiver in the coherent radar is a kind of phase-locked loop and is essentially a linear filter. As such, it can be made optimum in a Wiener sense, and, because it tracks, can be made very narrow band. However, it has a threshold below which it loses lock and a minimum bandwidth determined by the self noise of the oscillator in the loop. Thus, there is a lower limit of signal level which can be tracked. Typically, the minimum bandwidth is a few cycles per second and the threshold occurs when the signal power is equal to the noise power in the loop bandwidth.

The coherent radar has been used with targets which return a noncoherent signal, e.g. the Moon and Venus, and it appears to track the center of the returned spectrum. However, the analysis of the tracking process with the odd statistics of these signals is difficult and not yet completed.

The noncoherent CW radar uses a receiver which is a kind of spectrum analyser. No attempt is made to track the returned signal, rather the signal

is processed to determine the power spectrum. The form of the radar is shown in Fig. 5 with a passive target. A standard frequency is transmitted to the target and reflected back to the receiver. In general, the returned signal is shifted in frequency by the Doppler effect and spread in frequency by the reflecting characteristics of the target. A special feature of the receiver is the programmed local oscillator. This is a device which uses an ephemeris to automatically tune the local oscillator to remove the Doppler shift from the returned signal. An ephemeris is essentially the best estimate of the path of the target to be tracked. In the case of a celestial body or a spacecraft at an extreme distance the ephemeris is usually quite good, at least good enough to

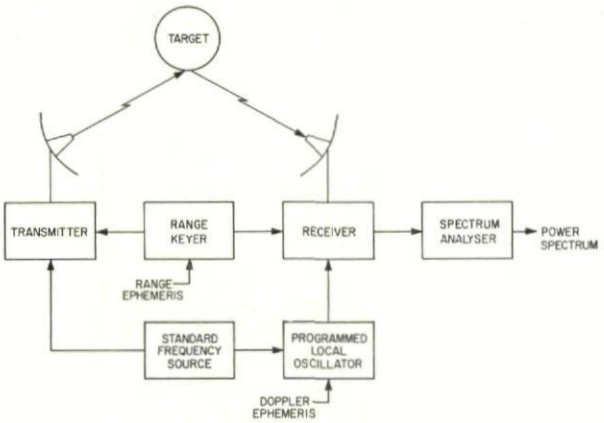


FIG. 5. *Non-coherent CW radar.*

tune the received signal into the range of the spectrum analyser and keep changes in frequency from blurring the spectrum while it is being measured.

The spectrum analyser itself is in two parts: a digital device which computes the autocorrelation function of the received signal continuously in real time and a small digital computer which computes the Fourier transform of the cumulative autocorrelation function whenever a power spectrum is desired.

Since the digital devices are not subject to drift, the spectrum may be determined for a signal of any desired duration. The ability to determine the spectrum of a returned signal over a long period of time is made useful by the fact that the spectrum analyser is arranged to overcome the effects of noise. This is accomplished by operating the radar as a Dicke radiometer. The transmitter is turned on and off for short intervals of time, typically 1 s. The receiver determines the spectrum of the returned signal plus noise when there is a returned signal and of noise only when there is no returned signal. The best estimate of the signal spectrum is then the difference of the two spectra. The receiver must, of course be told when to expect a signal. This is the function of the device labelled range keyer in Fig. 5. Note that

errors in the range ephemeris cause a loss in signal power about 27/1000 mi of range error when a 1 s on and off period is used. Ordinarily the distance to the celestial bodies and to spacecraft is known to within a few thousand miles.

The noncoherent CW radar does not suffer from the limitations given above for coherent CW radars. Signals have been detected at least 15 db below the threshold of the best coherent receivers known. In addition, the spectra from certain bodies, e.g. Venus, are not symmetrical. There are often prominent features on one side or the other which persist for several days and move slowly across the spectrum, apparently as the planet rotates. The presence of such an asymmetry certainly introduces a bias into the Doppler measurement as made by a phaselocked receiver. With the spectrum analyser, the location of the central peak of the spectrum can usually be determined.

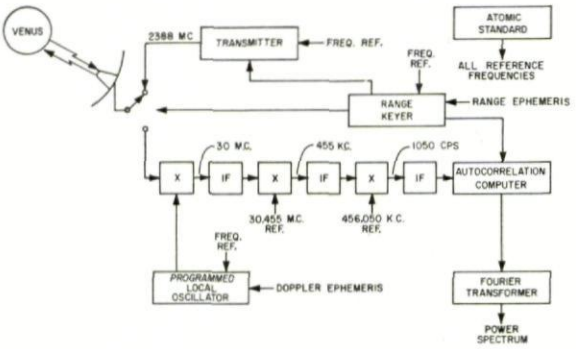


FIG. 6. 1964 *Venus radar system.*

Where the central peak is distorted by a feature close to the center, the shape of the spectrum shows this and the velocity measurement from the spectrum can be discarded.

A more detailed diagram of the radar used in the JPL 1964 Venus experiment is shown in Fig. 6. Note that a single antenna is used in this radar, not by diplexing, but by time sharing. Since the time required for a signal to propagate to Venus and back is a number of minutes, the antenna is used for transmitting for one round trip followed by one round trip time for receiving. The switchover time is a small portion of the round trip time. The spectrum analyser is stopped during the transmit portion of the cycle. but the information is not destroyed. The nature of the spectrum measurement allows the use of noncontiguous pieces of signal. Being able to use one antenna is a distinct advantage.

In the system of Fig. 6, the frequency source is an atomic oscillator with an accuracy of a few parts in 10^{11}, and the programmed local oscillator is capable of tuning to within 14 c/s at the received frequency.

Figures 7 and 8 give an indication of the range in signal level which may be handled by the radar. Figure 7 is a spectrum of the return from Venus

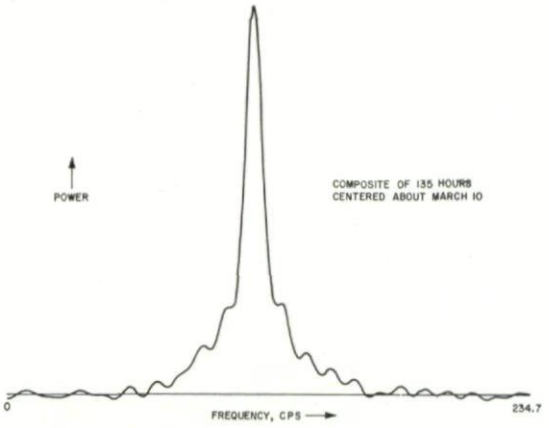


FIG. 7. *Venus power spectrum—long duration.*

using 135 hr of signal. Figure 8 is a spectrum taken when Venus was much closer and using only $\frac{1}{2}$ hr of signal. The resolution of the spectrum can be adjusted by adjusting the sampling rate of the device which computes the autocorrelation function. Since an increased resolution requires a slower sampling rate, the limitation on the resolution is set by the local oscillator rather than the spectrum analyser and a $\frac{1}{4}$ c/s is the limit achieved. This implies that the radar is able to distinguish increments in velocity of 1.6 cm/s. However, this requires that the ephemeris be smooth to this amount which poses some special problems in computing.

All in all, it is probably fair to say that the accuracy with which radar can measure velocity is about 6 cm/s. However, the measurement is unbiased so that a number of measurements can be smoothed together, for example by fitting an orbit.

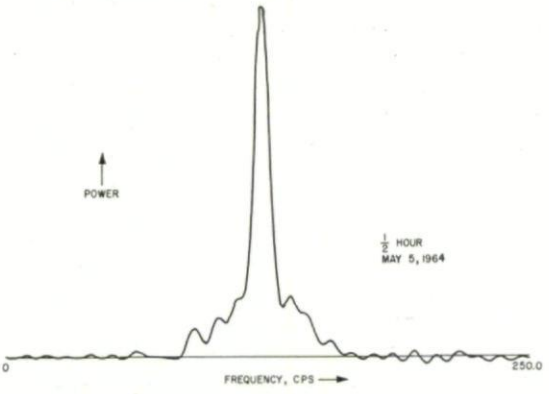


FIG. 8. *Venus power spectrum—short duration.*

4. THEORY OF PSEUDO-NOISE RANGING

A radar determines range by measuring the time required for a signal to propagate to the target and return. This may be done in a CW radar by modulating the transmitted signal with a suitable periodic waveform, recovering the modulation from the returned signal, and measuring the phase of the returned modulation relative to the phase of the transmitted modulation. This phase difference is a measure of the round trip propagation time and, therefore, of the range. Because the carrier is continuous, the modulation may be continuous; the receiver can be arranged to track the modulation, and the phase measurement can be made continuously to yield a continuous determination of range. Figure 9 is a simplified diagram of such

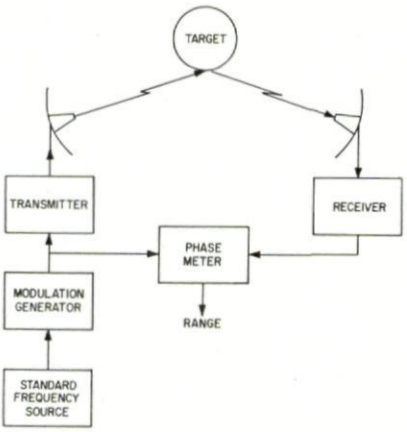


FIG. 9. *CW radar ranging system.*

a system. Note that the modulation waveform is timed by a frequency standard. This assures that the phase measurement, which is in terms of the period of the transmitted modulation, is an accurate representation of the round trip time.

There are several characteristics required of the modulation waveform. One is that it be continuous and periodic with a period long compared to the round trip time or at least compared to the *a priori* uncertainty in the round trip time. A second, is that it must be structured so that it is possible to measure the phase very accurately in order to achieve high range resolution.

These characteristics can be obtained in a binary waveform derived from a pseudo-random sequence. A pseudo-random sequence is as equence of binary digits which has certain random properties. The sequence is converted to a waveform by assigning a period of time called a digit period to each digit and causing the waveform to have a unit positive value when the corresponding digit is a ZERO and a unit negative value when the corresponding digit is a ONE. The sequence is repeated to make an indefinitely long periodic waveform called a *ranging code.*

The random properties of importance to ranging can be described by the autocorrelation function of the waveform. An example of a pseudo-random sequence, the corresponding waveform, and the autocorrelation function of the waveform are shown in Fig. 10. The interesting feature of the autocorrelation function is that, for all the values of τ which differ from a multiple of the period by more than one digit period, the value of the autocorrelation function is uniformly low. This implies that the phase of the returned modulation in the ranging system can be determined to within one digit period of the ranging code by crosscorrelation of the returned modulation with a local model of the code. Since a digit period of 1 ms is typically used, the phase can be determined to within 1 ms or the range to within 150 m. (Ways will be described later to increase the resolution without decreasing the digit period.)

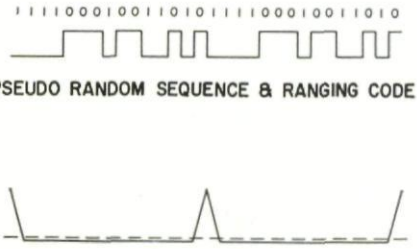


FIG. 10. *Ranging code.*

It has been tacitly assumed that the wave period of the ranging code is of sufficient length to resolve the *a priori* uncertainty in range. It turns out that this is no problem, since it is possible to generate codes whose periods are many hours long even with 1 ms digit periods. Therefore, a ranging code derived from a pseudo-random sequence fulfills our first two requirements: it can have a period as long as needed and the phase can be measured to within one digit period irrespective of the period of the code.

It turns out that there are some additional requirements. Since the systems under discussion measure very long ranges, the returned signal is weak and the individual digits in the code cannot be recovered from the noise. The phase of the returned signal is determined by a series of trial crosscorrelations between the returned code and a locally generated model of the code. If the local model has the same phase as the returned code, there is a correlation which is detected. If not, the lack of correlation is also detected, but no information is gained as to what the correct phase might be. If the code is very long, a great many trial correlations might have to be made before finding the correct phase. For example, if a code derived from a sequence of length one million were used and there were no *a priori* information about the phase of the returned code, a half-million trial correlations on the average would be required to find the correct phase. This number can be significantly reduced by forming the required long ranging code as a combination of a number of shorter codes. If several continuously repeating

P_1 = 23 μ SEC.
P_2 = 31 μ SEC.
P_3 = 47 μ SEC.
P_4 = 103 μ SEC.
P_5 = 127 μ SEC.
$P_1 + P_2 + P_3 + P_4 + P_5 = 331$
$P_1 \times P_2 \times P_3 \times P_4 \times P_5 = 438,357,391$
MAXIMUM RANGE = 65×10^6 km

FIG. 11. *Components in a ranging code.*

pseudo-random sequences with relatively prime periods are combined digit by digit according to any nontrivial Boolean function, the period of the resulting combined sequence is equal to the product of the periods of the component sequences. Furthermore, if the correct Boolean function is chosen for combining the sequences, the crosscorrelation function of the code derived from the combined sequence and a code derived from any one of the component sequences is periodic with the period of the component, has a peak when the component code is in phase with the component in the combined code, and has a uniformly low value when they are out of phase by more than one digit period. If such a combined code is used for ranging, the phase of each component can be determined by cross-correlating with the component code. Since the phase of the combined code is uniquely determined by the phases of the components, a number of trial correlations equal to the sum of the lengths of the component sequences need be performed in order to select the correct phase from among a number equal to the product of the lengths of the component sequences. An example of components which might be used in a ranging code is shown in Fig. 11 where the p_i are the periods of the several components.

The final requirement of the ranging code is that it must be capable of being tracked by the receiver. This is accomplished by including a component in the code which is called a clock. A clock is a code derived from a sequence of alternating ZEROS and ONES. The tracking portion of the receiver is shown in simplified form in Fig. 12. The received signal is a ranging code multiplied by a clock. The two signals have the same digit period so that the product is also a binary waveform with the same digit period. If the code from the code generator is the same as the incoming code

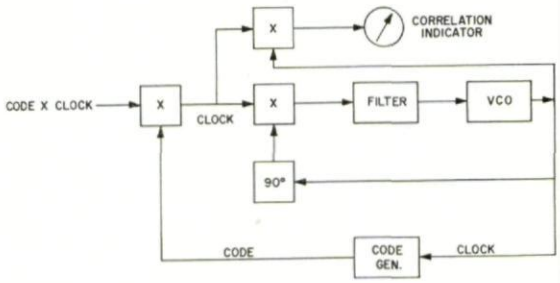


FIG. 12. *Double loop code tracking system.*

and has the same phase, the output of the first multiplier will be a clock. The inner loop is a phase-locked loop which locks to a clock and whose output times the code generator. If the phase of the incoming signal tends to change (the inner loop tracks the change in the phase of the clock and causes the outer loop to follow the phase of the code.

Thus, the double loop tracks the incoming code signal. If the code from the code generator is not in phase with the incoming code, the double loop system is used to acquire the incoming code, i.e. determine its phase by a series of trail crosscorrelations. The incoming code is structured so that it has a clock component which can be immediately acquired by the inner loop.[5,6] This structure assures that the phase of the code from the code generator will stay constant relative to the phase of the incoming code and, in particular, that their digit periods coincide. The code generator may then be manipulated so that its output is one of the components and the several phases of the component are produced one after the other. The first multiplier has a clock output whenever the two codes are the same, or more generally whenever the two codes are correlated.

Thus, the clock output is larger when the component code from the local code generator is in phase with the component in the incoming code than when it is not. The amplitude of the clock component and, therefore, the value of the crosscorrelation, is determined by the upper multiplier (synchronous detector). Since the several phases of the component code are almost orthogonal, the theory developed for block telemetry codes applies to the acquisition procedure, and the time required for acquisition can be computed along with the corresponding probability of error.[7,8,9] If automatic machinery is used to conduct the acquisition instead of a person reading the meter and controlling the coder, the acquisition can be done very quickly, typically a few seconds or tens of seconds for a radar designed to operate out to lunar distances. The theory has been fully checked by running a million and a half acquisitions under carefully controlled laboratory conditions.[10]

5. RANGE MEASUREMENT WITH A COHERENT CW RADAR

The basic range measuring system described in the last section was applied to a bistatic radar used to track the Echo I and the Courier satellites in the latter part of 1960 and the early part of 1961 primarily to test the fundamental concepts. Although the system did not contain all of the refinements later developed, for example automatic machine control of the acquisition process, all the basic aspects of the technique were used and the correctness of the whole approach was demonstrated. Useful tracking data were obtained on three consecutive passes of the Courier satellite on January 31, 1961. Angle and Doppler data were obtained from all three passes, range data only from the first and third.

To evaluate the ranging data, an orbit was computed from the angle and Doppler data of the three passes. The range was computed from the orbit and compared to the measured range. The range measured varied from

1,900,000 to 2,400,000 m with an average discrepancy between the range from the orbit and the measured range of about 20 m. This discrepancy includes the effect of all the inaccuracies in the measurement of angle and Doppler as well as all of the shortcomings in the orbit computation process. The good results obtained from this experiment together with a careful analysis of the probable sources of error led to the conclusion that if the technique were applied to spacecraft the range could be measured to within 15 m at lunar distances and to within 50 m at planetary distances. Accordingly, development of equipment and techniques was continued to produce a spacecraft system which is to have its first application in the *Mariner C* mission in late 1964.

The ranging system for spacecraft makes use of the same basic radar and transponder that have already been described in connection with the Doppler measurements. There are two types of ranging channels which may be used in the transponder, one for short ranges (out to lunar distances), the other for long ranges (out to planetary distances). In the first type (Fig. 13)

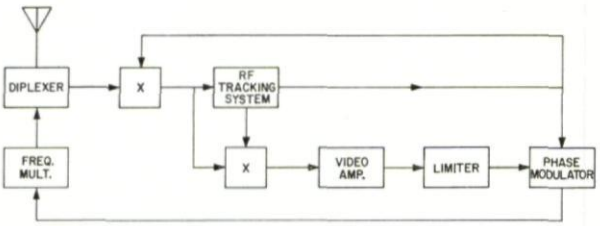


FIG. 13. *Turnaround transponder.*

called a turnaround transponder, the ranging modulation is synchronously detected, amplified by a video amplifier, hard limited to maintain a constant signal level from the video amplifier even when only noise is present, and used to modulate the returned carrier. This is a particularly simple way to handle the transponding of the ranging signal, since it requires no manipulation of the codes or determination of their phases in the spacecraft. The limitation of this approach is that the signal used to modulate the returned carrier has noise in it from a RF bandwidth of approximately 3 mc/s. As the signal level into the transponder decreases, the signal to noise ratio at the output of the limiter decreases. The presence of noise at the output of the limiter is not harmful in itself, but, because the limiter output has constant power, it reduces the power in the ranging modulation on the returned carrier. When the modulation power on the returned carrier becomes too low for the ground receiver to acquire and track the system will not operate. Normally, this occurs long before the noise modulation on the return carrier causes any effect because the bandwidth of the noise is very wide compared to the bandwidth of the ground receiver.

The ground receiver used with the transponder is shown in Fig. 14. The particular frequencies given are those for the NASA/JPL DSIF, but the technique is not dependent on the frequencies used. There are several

aspects of this receiver which have not been discussed. First, the double loop code tracking system is somewhat different from that shown in the previous section. The overall form is the same and functionally the loop performs the same operations but the first multiplier is a mixer at the second IF frequency rather than a multiplier at DC. The received signal goes into this mixer as a phase modulated signal at 10 mc/s. The local code balance modulates the 10 mc/s reference to the mixer. The mixer performs two operations simultaneously; the 10 mc/s signal is synchronously detected and the

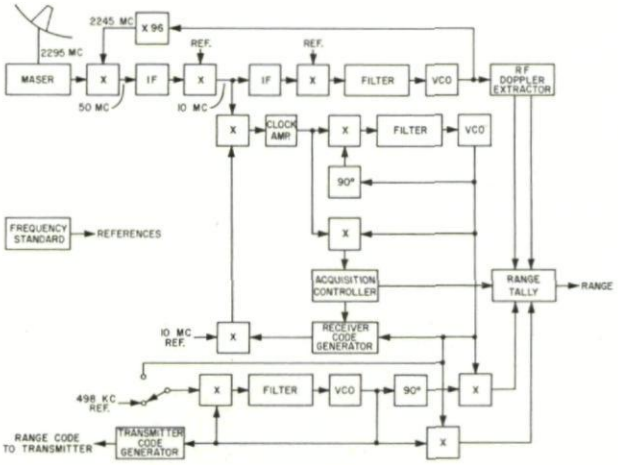


FIG. 14. *Ranging receiver.*

received code is multiplied by the local code. The output of the mixer is a clock whose amplitude is proportional to the correlation between the two codes. This technique has the very great advantage of not requiring a multiplier with a wide band input and DC output. Since the mixing is done at low signal levels, problems of linearity and noise handling are not severe. The clock amplifier is used to provide additional gain comparable to that of the second IF amplifier in the main receiver channel. However, the amplifier is narrow band so that the phase and amplitude detectors which follow it are not subjected to wide band noise. Their outputs are DC but their signal levels are high enough to overcome noise.[11]

The clock switch loop is part of the mechanism for measuring the phase between the codes. It allows the timing for the transmitter code (the code used to modulate the transmitter) to be derived either from the master frequency standard or from the voltage controlled oscillator (VCO) in the clock tracking loop. When the timing is derived from the VCO, both the transmitter code and the receiver code (the code which is used in the code tracking loop) have the same timing and may be synchronized. At this time, the phase between the codes is zero and the phase counting device, called the range tally, is set to zero.

If the switch loop is changed to obtain its input from the frequency standard, the timing for the two codes will, in general, be different and a phase difference will accumulate. The clock Doppler system is used to determine this phase difference. For each two-digit periods of phase difference between the codes the phase difference of the respective clocks is 1 c/s. For each two-digit periods of phase difference between the two codes the output of each clock Doppler detector varies through 1 c/s of a sinusoid. Since there are two clock Doppler detectors the direction of the phase difference can be determined and the range tally keeps a running sum of the phase difference between the codes. Once the receiver has acquired the clock, the change in phase is because of the two-way Doppler, hence the name Doppler detector. During the acquisition process the several components of the receiver code are stepped in phase by some integral number of digit periods to bring them into phase with the components of the received code. All of these changes in phase are accumulated by the range tally while it also keeps track of any changes due to Doppler. When the received code has been acquired, the receiver code is in phase with it and the range tally has an accumulation of the net phase difference between the transmitter code and the receiver code and hance a total of the round trip propagation time.[12,13] As the range changes, the range tally is kept up to date by the inputs from the clock Doppler detectors.

The clock Doppler detectors are able to provide an input to the range tally whenever the round trip path changes by one quarter of a clock wavelength which is equivalent to a change in range of 75 m for a 1 μs digit period. However, the clock tracking loop is able to follow the average phase of the clock with an accuracy of a few meters. To achieve a resolution which is commensurate with the accuracy, use is made of the fact that the transmitted code is coherent with the transmitted carrier, i.e. both are derived from the same oscillator. This means that there is a fixed relationship between the Doppler on the returned carrier and the Doppler on the returned code. For each cycle of clock Doppler there are some fixed number of RF Doppler cycles. Once the received code has been acquired, the range tally increases the resolution of the range measurement by counting RF Doppler cycles instead of clock Doppler cycles. Ordinarily the resolution which would be obtained from the Doppler on the carrier itself is not required so the Doppler on some IF is used instead. In the system shown in Fig. 14, the Doppler on the IF is used instead. In the system shown in Fig. 14, the Doppler on the IF at 143.4 mc/s is used to give a resolution of 1.04 m.

The system just described is to be used on *Mariner C*. Both the transponder and the ground system have undergone extensive tests in the laboratory and in simulated flights, using helicopters. The specifications for the system call for an overall accuracy of the range measurement of 15 m. but laboratory and helicopter tests achieved an accuracy of better than 2 m.[14]

The capability of the system can be extended to longer ranges by reducing the bandwidth of the ranging channel in the transponder. This is done by incorporating a full ranging receiver complete with coders in the transponder as shown in Fig. 15. Such a transponder is called a *planetary transponder* since

it is designed to operate at planetary distances. The advantages of the planetary transponder is that it has a very narrow noise bandwidth in the ranging channel, typically 1 or 2 c/s in contrast to the 3 mc/s for the turnaround transponder. This is because the only noise which goes through the channel is that which passes the clock tracking loop in the double loop system. The disadvantage is that the code must be acquired in the transponder as well as on the ground. The acquisition procedure is involved, but makes use of the fact that if the code is acquired on the ground first, the *a priori* knowledge of the range allows the determination of the phase of the codes in the transponder. The phase of the code which modulates the ground

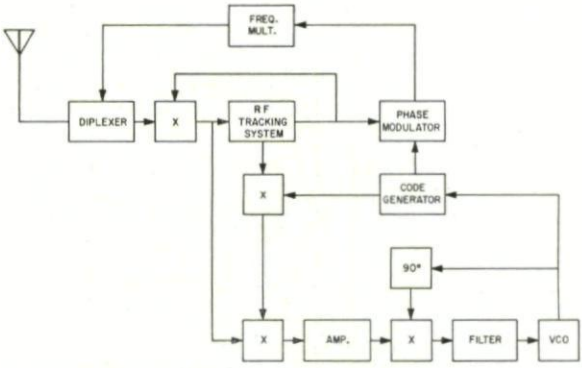


FIG. 15. *Planetary transponder.*

transmitter can then be adjusted so that when the code arrives at the transponder, it will have about the correct phase. If the phase of the ground transmitter code is then slowly varied, the transponder code tracking system locks when the code arriving at the transponder passes through the correct phase. The disadvantage is primarily in the increased manipulation required and, of course, the increased complexity in the transponder. The requirement for *a priori* knowledge of the range does not ordinarily pose a problem because once a spacecraft has been tracked to planetary distances, its position is quite well known and the purpose of day-to-day tracking is only to refine that knowledge.

The accuracy of the planetary transponder system is not quite as good as the accuracy of the turnaround transponder system. This is not because of the greater distances involved (the accuracy is independent of the distance), but because there are two tracking loops in the round trip path instead of one. Also, a loop aboard the spacecraft cannot be built as well or kept in adjustment as precisely as one on the ground. Careful analysis backed up by extensive laboratory testing indicates that 50 m is a conservative estimate of the accuracy of measurements and present radar system capability would allow such measurements to be made to anywhere in the solar system at least out to the orbit of Mars.

6. RANGE MEASUREMENT WITH A NONCOHERENT CW RADAR

The methods already described can be extended to measure range with a noncoherent CW radar. As discussed, the noncoherent radar receiver does not track the returned carrier but relies on a programed local oscillator to continuously tune the receiver and, in the monstatic case, a programed range keyer to control the time sharing of the antenna between the transmitter and the receiver.

One of the simplest ways of measuring range with such a radar is shown in Fig. 16. The transmitter reference signal is phase modulated by a ranging

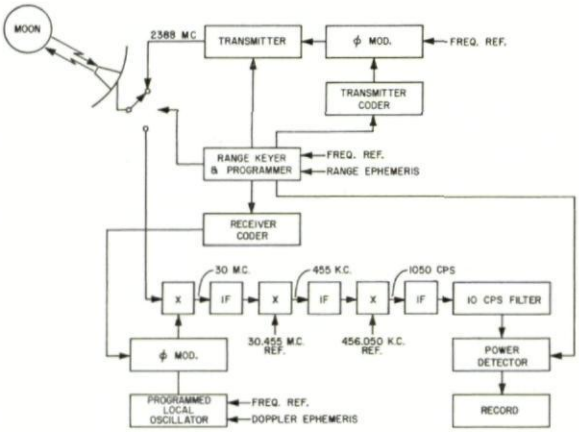


FIG. 16. *Open loop non-coherent ranging system.*

code to an angle such that, after the frequency multiplication in the transmitter, the carrier is balanced modulated. Since the receiver does not track the carrier no power need be reserved for it; all of the power may be put into the modulation. Similarly the LO signal which is applied to the first mixer in the receiver is also balanced modulated by a ranging code that is programed to have the proper phase angel relative to the transmitter code according to the range ephemeris. If the target were a *point* target and both the range and Doppler ephemerides were correct the output of the first mixer in the receiver would be an unmodulated carrier at the IF frequency. After amplification and heterodyning, the signal is filtered through a narrow band filter and power detected. The power detector includes an integrator and the average power is recorded. In the case or a large target such as the Moon, the returned signal may be considered to be composed of many components, each with a phase difference in the modulation proportional to the distance it has traveled, and a Doppler shift on the whole signal due both to the motion of the entire target (Moon) relative to the tracking station and to the rotation of the target as seen from the tracking station. The programed local oscillator removes the Doppler due to the relative

motion of the entire target. If the programed receiver code is adjusted to have the phase of the code returned from the nearest part of the Moon (the front cap), that code will cause an output from the first mixer which is proportional to the power in that portion of the signal and which is at the first IF frequency. A signal which returns from any other part of the target will, because of the correlation properties of the code, cause an output from the first mixer which is at least as wide band as the original modulated signal and which has a Doppler shift in addition. The recovered signal from the front cap will pass through the narrow band filter and be detected, signals from other parts will not.

In the discussion so far, it has been assumed that the range ephemeris is perfectly correct in which case there is no need to make a measurement. Of course, it is not perfectly correct. To obtain useful data the ephemeris is assumed to be correct and the power returned from the target at the ephemeris distance, with an appropriate integration time, is measured. The radar is operated as a Dicke radiometer by means of switching not shown in the figure. The range programer is then adjusted to cause the receiver code to be offset from the ephemeris position by one digit period and the run repeated. In our case when operating on the Moon the ephemeris turns out to be a little short so that no power is measured on the first run and the adjustment is made for the second run to offset the receiver code so as to increase the phase difference with respect to the transmitter code. This process is continued until the front cap is reached and usually a number of steps beyond. Each step beyond measures the power from what is called a range zone, which is a ring centered on the front cap.

The system which has been used at JPL on the Moon uses a 1 μs digit period so that each range zone is 150 m deep. The front cap is approximately 40 km in dia. or 1250 sq km. By the properties of a sphere all range zones have the same area.[15] There seems to be no doubt that the measurements are at least that accurate. The problem of using such measurements to improve the ephemeris is severe because of the local relief, which may be several km, and the irregularity of the figure of the Moon.

A similar system has been used to track the planet Venus. Since Venus is much farther away than the Moon, the range gate was widened to recover more power by admitting returns from a larger area. The digit period most used was 125 μs which gives a range zone of 1.875 km depth. Furthermore, the system was arranged to track in range. Note that, because of the *a priori* knowledge of range, a multiple component code is not required and no clock is needed for the open loop system used on the Moon. The tracking system used on Venus also dispenses with the clock and makes use of the shape of the scattering law for reflections from the planet. If the range gate were moved continuously across the front face of the planet, it would be observed that the power measured by the power detector in Fig. 15 would first increase as more and more of the range zone was on the planet but as soon as the entire range zone was on the planet further movement would cause the power to decrease because the normal to the area from which the signal was being returned would be at an increasing angle to the line of

sight. Thus, if there were two adjacent range zones, they could be adjusted to have equal returned power by having the one partially in front of the front cap and the other into the planet far enough to be past the peak of the reflection characteristic. Such a pair of range zones could be made to track the planet by comparing the power and causing both zones to move out of the planet if the power in the first zone is greater than in the second and vice versa. This is the scheme that is used except that the two measurements are not made simultaneously. The power is measured first in one zone and then in the other. A comparison is made, and, if the power is not the same, the phase of the receiver code is adjusted. It is not necessary to operate the receiver as a radiometer since the noise is automatically subtracted out in the comparison process. The block diagram for the system is the same as that in Fig. 15, except that the output of the power detector is fed back into the range programmer which includes a stored program machine capable of performing the necessary computations and controlling the receiver coder accordingly. Since the antenna is time shared, it is not available to the receiver half the time. During these intervals the range rate is extracted from the ephemeris to continuously adjust the phase of the receiver coder. During the time that the receiver is available the ephemeris rate is still used to help adjust the phase of the receiver coder.

The system just described was used during May, June, and July of 1964 to track the planet Venus.[16] Typical runs were for 2 hr of which half the time was receive time. The noise on the range measurement was typically 5 km peak to peak. The residuals of the range compared to the ephemeris over a 12-day period were used to determine a straight line. The rms variation of the residuals about the line was 2 km.

7. CONCLUSIONS

A number of methods by which a CW radar may be used to measure radial velocity and range have been presented. Experience has shown that, if the returned signal is coherent and strong enough to be tracked by a phase-locked loop receiver, a CW radar can determine the radial velocity to a fraction of 1 cm/s. If the returned signal is not coherent or the signal is not strong enough to be tracked by a phase-locked loop receiver, but there is a reasonably good ephemeris for the target, the radial velocity can be determined to a few centimeters per second. The noncoherent method has so far been applied only to passive targets but could certainly be applied to spacecraft.

Range measuring experiments with passive targets combined with extensive laboratory and helicopter experiments indicate that accuracies of 15 m can be achieved if the returned signal is coherent and the target is either small and passive or uses a simple turnaround transponder. Of course, the signal must be strong enough so that the carried can be tracked by a phase-locked loop receiver. If the distance is great enough so that a planetary transponder with a range tracking loop is required, an accuracy of 50 m should be achievable.

Techniques for measuring range when the returned signal is noncoherent have so far been applied only to celestial bodies. The accuracies achieved, 150 m for the Moon and 2 km for Venus, were limited primarily by the roughness of the target and the digit period needed to obtain the necessary reflecting surface on the target. Certainly higher accuracy could be achieved with a small target or with transponding spacecraft.

The several methods which have been described have all been mechanized and tried by actual experiment. They represent the present state of the art and show that it is, in fact, well developed. However, none of the techniques has yet run into any fundamental limitation, such as variation in propagation velocity or opacity in space, which limits its accuracy or the distance over which it may be used. It is, therefore, safe to say that the techniques presented will be developed further and new ones found.

REFERENCES

1. Stevens, R. and Victor, W. K., The Goldstone station communications and tracking system for Project Echo, *Technical Report No.* 32-59, Jet Propulsion Laboratory, Pasadena, 1960.
2. Victor, W. K. and Stevens, R., The role of the jet propulsion laboratory in Project Echo, *Trans. IRE*, **SET-7,** 20–28, 1961.
3. Easterling, M., A skin tracking radar experiment involving the Courier satellite *Trans. IRE*, **SET-8,** 76–84, 1962.
4. Mathison, R. P., Tracking techniques for interplanetary spacecraft, *Technical Report No.* 32–284, Jet Propulsion Laboratory, Pasadena, 1962.
5. Easterling, M., The evolution of ranging codes, *Research Summary*, 36-10, **1,** Jet Propulsion Laboratory, Pasadena, 26–27, 1961.
6. Titsworth, R. C., Optical ranging codes, *Trans. IEEE*, **SET-10,** 19–30, 1964.
7. Easterling, M., Maximum likelihood acquisition of a ranging code, *Research Summary* 36–12, **1,** Jet Propulsion Laboratory, Pasadena, 81–84, 1961.
8. Golomb, S. W. (Editor), *Digital Communications*, Prentice-Hall (in preparation).
9. Viterbi, A., On coded phase-coherent communications, *Trans. IRE* **SET-7,** 1, 3–14, 1961.
10. Helicopter tests simulating lunar ranging, *Space Programs Summary*, 37–18, **III,** Jet Propulsion Laboratory, Pasadena, 50–53, 1962.
11. Easterling, M., Demodulation in ranging receivers, *Research Summary* 36–7, **1,** Jet Propulsion Laboratory, Pasadena, 62–65, 1961.
12. Ranging subsystem, *Research Summary* 36–2, **1,** Pt 1 Jet Propulsion Laboratory, Pasadena, 31–36, 1960.
13. Range tally, *Research Summary*, 36–3, **1,** Pt. 1, Jet Propulsion Laboratory, Pasadena, 47–50, 1960.
14. Helicopter tests simulating lunar ranging, *Space Programs Summary* 37–19, **III,** Jet Propulsion Laboratory, Pasadena, 28–29, 1962.
15. A range gated lunar radar experiment, *Space Programs Summary* 37–25, **III,** Jet Propulsion Laboratory, Pasadena, 38–47, 1963.
16. A precision long-range tracking radar system for planetary ranging, *Space Programs Summary*, 37–28, **III,** Jet Propulsion Laboratory, Pasadena, 1964.

CHAPTER 19

A DESCRIPTION OF A HIGH-PERFORMANCE MICROWAVE EXPERIMENTAL FACILITY

H. G. Weiss
Lincoln Laboratory, Massachusetts Institute of Technology,
Lexington, Massachusetts, U.S.A.

A versatile new ground station has been developed for experimental space communications, radar and radio physics investigations. This installation, which has been named HAYSTACK, employs, a precise 120-foot-diameter Cassegrain antenna with a surface tolerance which should permit very efficient operation at wavelengths of 3 centimeters and a useful capability a wavelengths as short as 8 millimeters. The antenna is protected from the environment by a metal spaceframe radome and employs a digital computer in its pointing control system. Through the use of a plug-in equipment box, which can be mounted behind the moving reflector, it will be possible to conveniently carry out a variety of experiments. A versatile transmitter power supply and a high-power microwave transmitter and a low-noise receiver are incorporated in the system.

1. INTRODUCTION

For more than a decade, the M.I.T. Lincoln Laboratory has been conducting exploratory investigations in the fields of communications, radar and radio physics. About five years ago, it became apparent that a new and more sensitive experimental facility would be required in the microwave portion of the radio spectrum to augment studies being conducted at longer wavelengths at the M.I.T. Millstone Hill Station in Westford, Massachusetts. The proposed investigations required a versatile installation capable of high-power transmission and low-noise reception in a variety of modes over the 1 to 35 Gc/s frequency range. This paper describes a new facility which has been recently established by the Lincoln Laboratory in Tyngsboro, Massachusetts, with support from the Electronic Systems Division of the U.S. Air Force. This new installation, which is commonly known by its code name HAYSTACK, is located approximately thirty miles northwest of Boston, about one mile from the Millstone Hill radar. A recent aerial photograph of this installation is shown in Fig. 1.

The integration of a number of novel concepts into this facility, plus the utilization of advanced transmitter, receiver, control, and data processing components has resulted in a very sensitive and versatile ground terminal

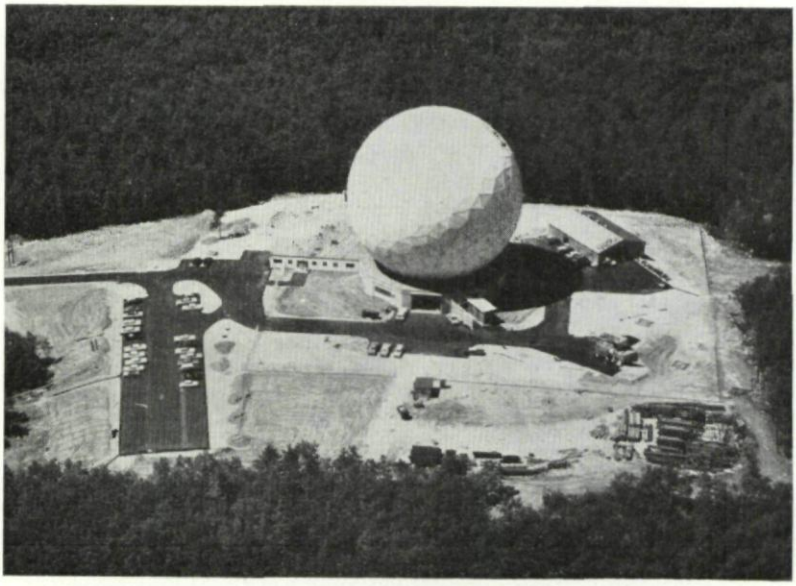


FIG. 1. *Aerial photograph of HAYSTACK Field Station.*

that will be used for space communications, radio astronomy, and long-range radar tracking experiments. A number of the significant aspects of the HAYSTACK program are:

1. The development of improved techniques for analyzing complex mechanical structures. As a result of this program, which has involved scale model tests, an extensive analytic effort, the development of improved quality control and measurement techniques, and structural tests on the (completed) antenna, it is now possible to predict the behaviour of large complex space frame and shell structures with an uncertainty of about one part in 100,000. This is about two orders of magnitude more precise than could be accomplished by using the design techniques available at the start of this program.
2. The uilization and evaluation of a metal-space-frame radome in a low-noise microwave system. This type of radome may be more generally employed in the future, since larger sizes are readily feasible, and they can be designed to provide good electrical performance over a wide frequency range.
3. The development of a hydrostatic bearing which will permit the positioning and control of massive structures with great precision. This class of bearing also offers promise of being essentially free from wear.
4. The use of Cassegrain reflector geometry together with the integration of a "plug-in" equipment room into the antenna structure permits the utilization of the antenna for a number of diverse applications and provides an efficient RF design that is compatible with the use of low-noise receivers and high-power microwave transmitters.

5. The integration of a digital computer into the real-time antenna control loop in a manner which provides great flexibility and permits a non-skilled user to utilize the antenna in an efficient manner.
6. The exploitation of recent advances in maser and high-power transmitter technology to provide increased system sensitivity for long-range communications and radar tracking experiments.

2. ANTENNA

2.1 *Performance Requirements*

Because of the wide variety of proposed uses for the HAYSTACK antenna, the specification of its operating frequency range was somewhat arbitrary. There were several specific radar and communication studies which required the use of a large aperture antenna at a wavelength of 3 centimeters, and other programs could utilize a large antenna at considerably shorter wavelengths. To be of practical value for the proposed communications and radar studies, the antenna efficiency would have to remain high for all pointing angles and throughout the range of environmental conditions that would be encountered in a New England site.

As discussed later in this report, it became evident early in the program that it would be unrealistic to undertake the development of a large, precison microwave antenna unless it were protected from the environment. The availability of a 150-foot-diameter radome influenced the decision to limit the diameter of the antenna to 120 feet, the largest that could be housed in a radome of this size. The longest nominal operating wavelength was chosen to be 30 centimeters, to permit the antenna to be used for hydrogen line and L-band radar studies.

A performance goal was selected which required an aperture efficiency only 0.3 db below that obtainable from an ideal 120-foot-diameter parabolic reflector at a wavelength of 3 centimeters. The fulfillment of this requirement would permit the antenna to be useful at wavelengths as short as 8 millimeters. This would represent a significant advance over previous antenna designs and would make the development program technically challenging.

The relationship between the surface tolerance of a reflector antenna and its aperture efficiency at different wavelengths is plotted in Fig. 2. To meet this specified performance criterion, the reflector surface should not deviate more than $\pm 0.025\lambda$ (rms) from an ideal reflector. For $\lambda = 3$ centimeters, this corresponds to ± 0.075 cm. In working with production drawings, manufacturing operations, and field measurement practices, it is more convenient to specify peak tolerances rather than rms tolerances, and since experience has shown that the peak surface deviation of a typical antenna is approximately 2.5 times its rms deviation, the ± 0.075 centimeter rms tolerance corresponds to a peak value of ± 0.187 centimeter or about ± 0.075 inch. Therefore, the antenna specification was written to require that all points on the reflector surface remain within ± 0.075 inch (peak-to-peak) from the ideal parabolic contour under the expected operating conditions. This allowable peak error of 0.075 inch must include all surface errors

resulting from gravity forces (varying with rotation), thermal effects, manufacturing errors, and field erection and calibration errors.

To meet this requirement, design and measurement techniques would have to be developed which would contribute an uncertainty of perhaps only 0.02 inch maximum. In a 120-foot-diameter structure, this corresponds to about 1 part in 70,000 and represents a substantially more stringent design requirement than has been achieved in any large movable structure to date.

A 120-foot-diameter antenna at a frequency of 10 Gc/s will produce a radiation pattern with a main lobe width that is 0.05° between 3-db response

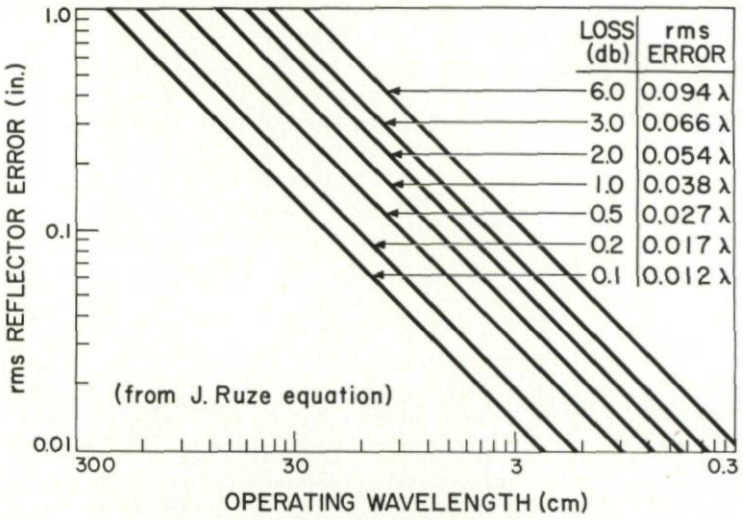


FIG. 2. *Loss due to reflector tolerance.*

points. For this reason, the specifications were written to require that the pointing error not exceed 0.005° (18 arc seconds) at slow tracking rates. While the tracking rates that would be encountered in most experiments would be relatively slow, it was expected that there would be occasions when it would be desirable to transfer from one object in the sky to another in a relatively brief period. For this reason, it was specified that the antenna be able to move from one pointing position to any arbitrary pointing position in the hemisphere and achieve a maximum pointing error of 0.005° in less than 60 seconds.

Because of the variety of experiments to be performed with the antenna, it seemed advisable to integrate a general purpose digital computer into the pointing control system. The use of an appropriate computer in this manner would enable the antenna to be directed from data obtained over telephone or radio circuits from a remote location, and rapid and accurate parallax and time-delay corrections could be introduced. The use of the computer would also enable semi-automatic compensation to be achieved for systematic pointing errors that might result from displacement of the feed with elevation

motion, lack of orthogonality in the elevation and azimuth axes, systematic errors in the shaft encoder, etc. It would provide the possibility of introducing corrections for the refractive index of the atmosphere, which could be significant at elevation angles below 15°.

A Cassegrain reflector system was selected because it combined good electromagnetic performance with a basic configuration that permitted the placement of electronic components near the center of gravity of the moving structure. The energy spillover from the primary feed horn of a Cassegrain antenna is largely in the direction of the "cold" sky and this is a desirable attribute for a low-temperature antenna. The size of the hyperbolic secondary reflector is chosen to be 9 feet 4 inches in diameter so that, at the lowest nominal wavelength, namely 30 centimeters, the size of the secondary reflector would equal that of the feed.

To provide a convenient way for carrying out a variety of experiments with the antenna, the concept of interchangeable "plug-in" instrumentation rooms was evolved. These "plug-in" rooms, mounted immediately behind the primary reflector surface, would contain low-noise receivers, radio frequency transmitter components and antenna feed systems. The Cassegrain geometry adopted for this antenna in conjunction with the "plug-in" room concept permits the use of low-noise receivers and high-power microwave transmitters, and requires only short waveguide runs without rotary joints. The "plug-in" room was specified to be 8 feet × 8 feet × 12 feet in size and have a maximum gross weight of 7000 pounds. An integral hoist system was to be incorporated into the antenna structure to facilitate the interchanging of "plug-in" rooms. Power, water, and other services would be conveyed to the equipment by cables with quick-disconnect plugs.

2.2 *Development Approach*

The techniques previously available for estimating the performance of an antenna required that the structure be approximated by a greatly simplified two-dimensional mathematical model, and study revealed that these simplifying assumptions would make it impossible to obtain a valid analysis of any complex, moving structure which was accurate to better than about 1 part in about 300.

Therefore, the specifications called for the development of a comprehensive analytic computer program to confirm the validity of any proposed design, the construction and testing of a $\frac{1}{15}$ structural scale model and the comparison of the performance of this model with that predicted by the computer program, preassembly of the complete antenna system at the contractor's plant, and then proof-loading and checking against predicted performance prior to shipment to the site. A further stipulation in the contract called for the submission of all drawings to Lincoln Laboratory for approval to release to production.

2.3 *Description of Antenna Configuration*

In December 1960, a contract was awarded by the Air Force to the North American Aviation, Inc., Columbus (Ohio) Division for the design,

construction, and installation of the HAYSTACK antenna system. The basic structural configuration proposed by NAA is shown in Fig. 3. It consists of a reflector backup structure comprised of five concentric tubular ring trusses interconnected by pretensioned diagonal rods. The main supporting ring, which is 60 feet in diameter, is bridged by two trunnion

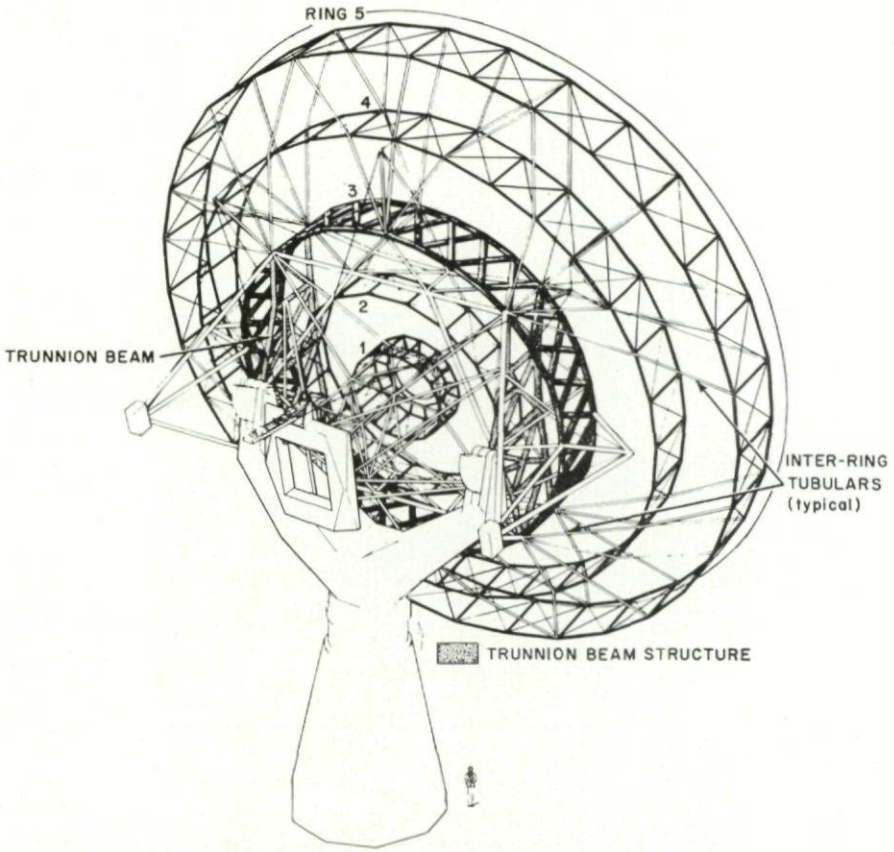


FIG. 3. *Basic structural configuration of HAYSTACK antenna.*

trusses that are attached to the main ring at its quadrant points and carry its load to the elevation bearings. Two ring trusses are positioned outboard of the main support ring, and two smaller ring trusses are positioned inboard of the main ring. The pretensioned interconnection rods are designed to work without stress reversals as the antenna rotates. In the trunnion beams and the support rings, where high compression loads occur, and where some stress reversals cannot be avoided, heavy-wall large-diameter tubing was used, and considerable effort was expended in the development of a bolted joint that was not susceptible to creep. This configuration results in a structure that is unusually stiff in proportion to its weight.

The reflector surface, which is supported from the backup structure by means of adjustable standoff studs, is made up of 32 inner panels and 64 outer panels. These panels consist of thin aluminium prestretched skins, 0.016 inch thick, bonded to a $\frac{1}{2}$-inch-thick aluminium honeycomb core. The panels are interlocked by shear keys. Adjustable expanders are located

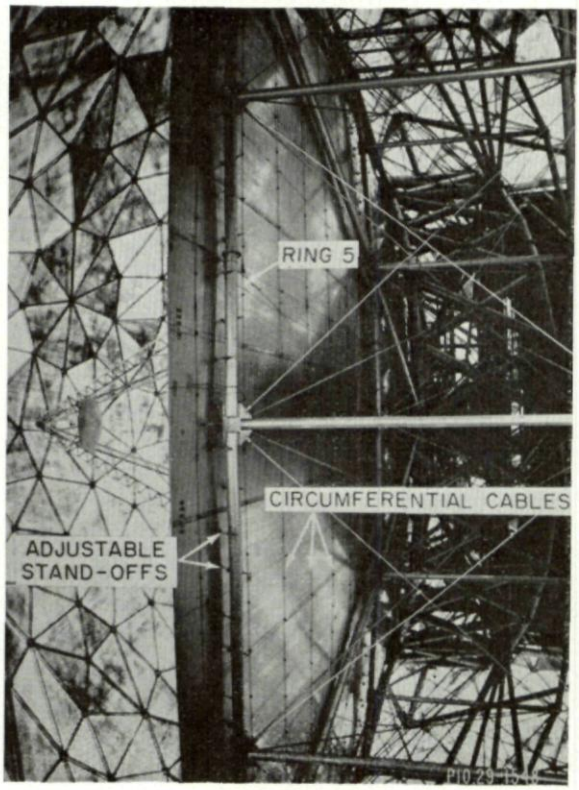


FIG. 3a. *Side view of antenna.*

at 2-foot intervals along the radial edges of each panel, to determine the gap between panels. One circumferential edge of each inner and outer panel is firmly attached to a 60-foot-diameter annular ring, which in turn is supported from adjustable stand-offs from the main ring. Each panel is also attached to two of the other ring trusses by adjustable stand-offs which are placed normal to the local contour. These have ball and socket joints at each end and thus take only axial loading.

The 96 individual reflector panels are made to act like a homogeneous shell by the use of 26 concentric circumferential cables located behind the reflector surface (Fig. 3a). These pretensioned cables, which act like large "elastic bands" are guided by adjustable rollers mounted on the back of the

reflector panels, close to the panel surface. The exact contour of the reflector surface can be adjusted by the interpanel expanders, the adjustable standoffs at the rings, and the spacing between the cable rollers and the back of the panel. By forcing the individual panels to work as a homogeneous shell, the stiffness of the reflector surface is increased about 10-fold, and the relatively lightweight reflector surface contributes significantly to the overall rigidity of the complete antenna system. When the reflector is rotated away from the vertical, the integral shell is restrained from lateral movement relative to the backup structure by 32 shear pins that are installed between the shell and the trunnion trusses.

The Cassegrain secondary reflector is a 9-foot 4-inch-diameter hyperbolic surface supported by a planar truss quadripod. The reflector size was selected to permit operation at wavelengths as long as 30 cm. The quadripod legs are attached to the main ring at the trunnion beam "hard" points an each truss is guyed by 4 pretensioned rods. The reflector surface was explosively formed from a 0.188-inch-thick sheet of aluminium which was then bonded to a 2-inch-thick honeycomb stiffener. The front surface of the aluminum sheet was then milled in a digital controlled boring mill. While the hyperbolic reflector axis was established in the boring mill, an optical flat, 2 inches in diameter, was carefully located at the center of the hyperbolic reflector with its axis coincident with the reflector axis to within 1 second of arc. In this manner, a relatively lightweight but very rigid precision secondary surface was obtained. The surface of the secondary reflector has a tolerance of ± 0.005 inch. The position and tilt angle of the secondary reflector are adjustable by lead screw actuators which are remotely controlled.

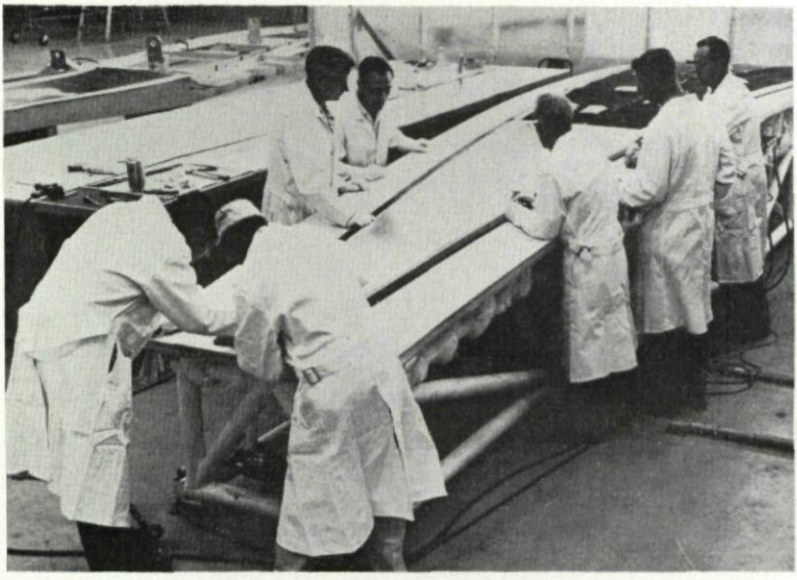


FIG. 4. *Reflector panel fabrication.*

2.4 *Quality Control and Measurement System*

Refined measurement and quality control techniques were developed by NAA to produce panels, which would conform to the theoretical contour with an rms deviation of approximately 0.010 inch when measured in a precision test fixture in a temperature-controlled environment. As shown in Fig. 4, these panels, which are nearly 30 feet long, were fabricated on a precision, doubly-curved mould and employed a thermal-setting epoxy to bond the prestretched aluminum skins to the honeycomb core. The panels were then placed in a precision drill and trim fixture, still in a constant-temperature room, where they were trimmed and 16 optical targets carefully located on each reflector panel surface. The tolerance in locating these targets within any given panel is 0.003 inch.

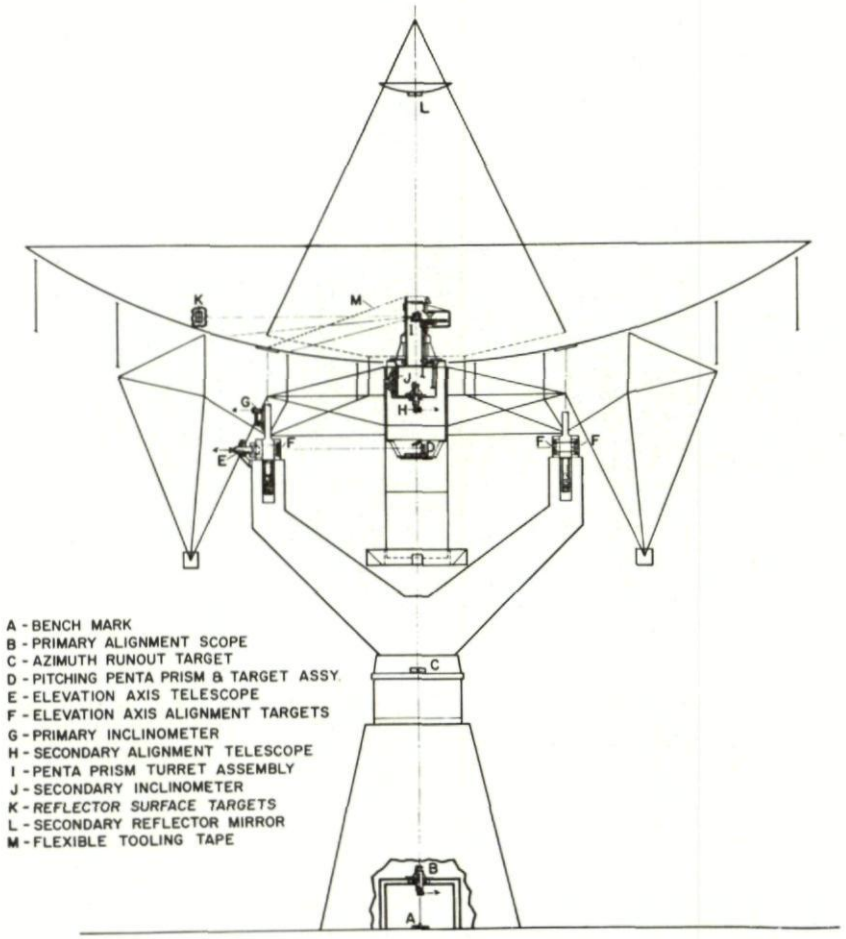


FIG. 5. *Optical alignment system.*

After the antenna has been assembled, the surface is adjusted by the use of the system of optical elements shown schematically in Fig. 5. The optical system consists primarily of alignment telescopes, precision levels, ruled circles, targets, pentaprisms, and a calibrated tape. Each element is the most precise unit of its kind available. To facilitate the use of the alignment system, the over-all antenna configuration has been designed to preserve optical paths from the ground through the azimuth and elevation axes.

FIG. 6. *Optical probe.*

The primary optical probe for adjusting the surface contour was built by Keuffel and Esser of Hoboken, New Jersey under subcontract from North American Aviation. It is shown in Fig. 6. This probe is mounted in a "plug-in" room, which has its center of gravity and total weight matched to that of a typical plug-in equipment room. A special chair on gimbals has been provided to permit an observer to use the optical probe while the antenna is rotated in elevation. The upper head of the calibration probe may be rotated and contains 8 indexable, fixed-angle pentaprisms that permit observation of 8 circular rows of targets on the reflector surface. Each pentaprism, which has been calibrated to have an error of less than 1 second

of arc, rotates the line of sight the proper amount to view a circle of targets on the reflector surface. This system does not require an outside reference, and it permits referencing all target points with respect to four "hard points" located where the secondary reflector quadripod legs join the main support ring. Experience in using the optical probe at the HAYSTACK site has indicated that, when the thermal environment is relatively stable (at night), the repeatability of the optical measurements on a night-to-night basis is within 2 seconds of arc. This corresponds to an uncertainty of about 0.007 nch in a direction normal to the line of sight at a radius of 60 feet.

2.5 *Bearing and Control System*

The bearings used for the elevation and azimuth axes of the antenna are of two distinctly different types. In the azimuth axis, an externally pressurized oil-film bearing is employed (Fig. 7). The lower portion of this bearing consists of a 6-foot-high load distribution cylinder which is located between the rotating portion of the antenna and the concrete tower. The azimuth bullgear, which is about 14 feet in diameter, is bolted to this distribution cylinder and serves as a support ring for the hydrostatic bearing pads. These pads, each of which has an effective surface of 8 $\times$ 14 inches, are attached to the support ring by bolts. Sixteen pads are used for the thrust bearing and 8 pads are used for the radial bearing. The combined weight of the pad, ring, and structural member is 42,000 pounds. Adjustable wedge block assemblies are used between the concrete tower and the load distribution cylinder to permit releveling of the bearing.

The upper load distribution ring is approximately 14 feet in diameter and 4 feet high. The ring, which weighs approximately 24,000 pounds, contains integral thrust and radial runner surfaces that have been flame-hardened and surface-ground. It is a very stiff weldment, which has been designed to distribute the concentrated loads that are conveyed from the yoke arms. The bearing pads operate with an oil-film thickness of approximately 0.005 inch. A temperature-compensated flow-control valve has been inserted between the pump, which is common to all pads in the bearing, and each bearing pad, to isolate the down stream effect from the other bearing pads The normal operating pool pressure is 250 psi, and an oil flow of approximately 26 gallons per minute is required for both thrust and radial bearings. This type of bearing should provide extremely smooth performance and have a long life.

Conventional self-aligning, spherical, roller bearing assemblies are used in the elevation axis. Because the diameter of the elevation bearings is relatively small (10.5 inches). the "stiction" torques are of very small magnitude. An oil circulation system is used to force-lubricate these bearings. The alignment of the bearing pads on the top yoke arms is accomplished by the use of positioning brackets and by shims under the pillow blocks. An optical path has been provided along the bearing axis to facilitate their alignment.

To insure that the antenna will not rotate beyond the nominal travel limits, a self-contained oil-air buffer stop has been provided. The buffers

HYDRAULIC DRIVE

POWER TRAIN			TACH DRIVE		
GEAR	NO TEETH	G.R.	GEAR	NO TEETH	G.R.
1	21	.266	7	96	4.57
2	79		8	21	
3	21	.250	9	105	4.38
4	84		10	24	
5	26	.046			
BULL	564				

DATA BOX DRIVE

CAM DRIVE			SYNCHRO DRIVE		
GEAR	NO TEETH	G.R.	GEAR	NO TEETH	G.R.
BULL	564	21.7	BULL	564	21.7
1	26		1	26	
2	30	.250	2	30	.250
3	120		3	120	
4	28	.199	4	28	.199
6	141		5	141	
7	64	.500	4	28	.199
8	128		6	141	

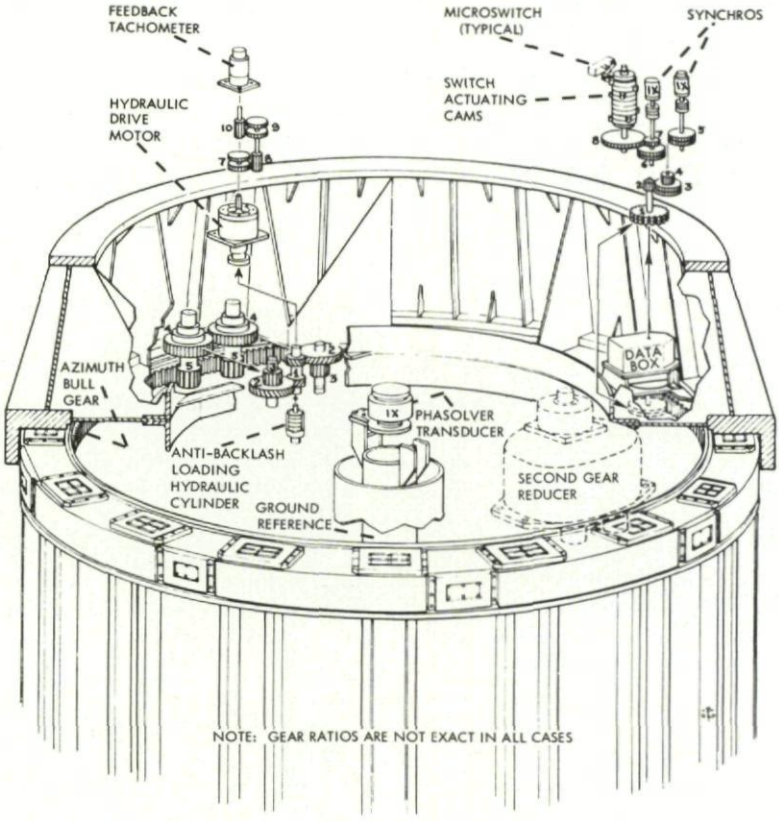


FIG. 7. *Azimuth bearing and gear train.*

are of the linear type, somewhat similar to those used in aircraft landing systems. The buffer systems have been designed to decelerate the antenna safely when it is moving with full drive torque and with maximum velocity.

The azimuth and elevation antenna drive elements each employ two hydraulic motors and gear reducer systems. Dual pinnion drives contact the bull gear at each gearbox and use two parallel gear trains hydromechanically loaded to minimize backlash (Fig. 7). The hydraulic servo motor, which

has a 20 h.p. rating, employs a series of radially placed pistons that push against a multilobe cam to make it rotate. This type of motor has extremely smooth lowspeed characteristics.

The HAYSTACK antenna has two angle-data systems. One is a conventional one-speed synchro-loop and provides a reduced accuracy standby position control for checkout operations and emergency control. The accuracy of the synchro readout system is $\pm 0.1°$. The more precise digital control and data system was specifically developed for the HAYSTACK antenna. A shaft-angle transducer consisting of two electrostatically coupled, 8-inch-diameter glass discs has been developed by the Telecomputing Corporation, LaMesa, California, for the program. One disc revolves on its own bearings and is coupled directly to the antenna shaft rotating with the antenna at one speed; the other disc is fixed to the encoder housing. Clearance between discs is approximately 0.003 inch. One disc contains three sets of conductive patterns consisting of paired 90° displaced sine waves. This electrostatic resolver provides 256 fine-electrical-phase vector rotations and 8 medium-phase vector rotations for one antenna shaft revolution. Quantizing electronics converts the electrical phase information into pulse form for use in a counter-comparator. The least significant bit of the precision shaft encoder corresponds to 2.47 seconds of arc. Specially designed shaft couplers are employed to allow small alignment errors between the encoder shaft and the moving antenna without causing angular errors.

2.6 *Analytic and Test Program*

In 1957, the Department of Civil Engineering at M.I.T. under the sponsorship of Lincoln Laboratory developed a computer program for use in structural analysis of large space truss structures. This program, capable of handling up to 60 joints, was modified and expanded to the extent that by 1960 a program was available that could analyze space structures with up to 4000 joints. The acronym, "STAIR" (STructural Analysis Interpretic Routine) was adopted for this program. The basic parameters of STAIR as programmed for HAYSTACK on the IBM 7094 are indicated in Fig. 8.

Basically, STAIR considers the HAYSTACK backup structure as comprised of truss members with pin-ended joints, but with no provisions made for member bending or torsion, or for moment transfer at the joints. A

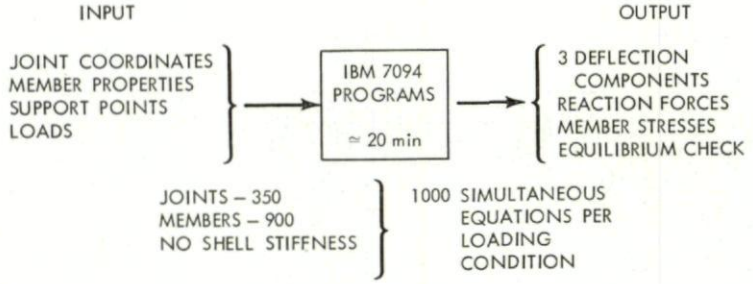


FIG. 8. *STAIR program.*

unique feature of STAIR is the method of preparing the input, which consists of dividing the joints into groups, which are called "units." For each unit a stiffness matrix is calculated, each element of which represents the resistance to displacement of a joint in the direction of one of the principal orthogonal axes. Successive steps of matrix addition and reduction are then carried out until the entire structure finally is represented in a single matrix. The final matrix is inverted and multiplied by the applied load matrix to obtain joint deflections. Deflections of the joints previously eliminated are obtained by back substitution. The program then calculates the force in each member from the deflections of its ends. An important feature of the program is a final status check, in which the forces at each joint are summed and tested for equilibrium.

With STAIR available in 1960, Lincoln Laboratory was able to check independently the analytical work submitted by North American Aviation. The original 2-dimensional hand calculations on the backup structure performed by NAA were not verified by STAIR. The analysis was in error by an order of magnitude (hand prediction 0.010 inch vs STAIR prediction 0.110 inch). STAIR, by virtue of its ability to look at local surface areas, was able to identify specific design deficiencies. The hand analysis could not do this, since it provides only maximum deflection data on an area basis and does not give point-by-point distortion information. This is an important distinction, since the HAYSTACK reflector panel contour is deliberately distorted, when it is adjusted in a face-up attitude, to minimize surface errors when the antenna is oriented to an attitude of 45 degrees.

A $\frac{1}{15}$-scale structural test model was constructed and tested for deflections under various loading conditions. STAIR was utilized to predict the model deflections with the applied loads in various tests. Correlation of model and computer results fell within 15 to 20 percent. This provided the first opportunity to establish the validity of the STAIR program.

By mid-1961, NAA had completed an independent analysis, using a separate NAA-generated computer program. There was excellent correlation between deflections computed by STAIR and NAA. Of the values tabulated, 93 percent agree within 0.001 inch, and the largest difference was 0.018 inch, at a point where the deflections had a value of 0.241 inch. It became quite evident with the agreement of both programs that redesign of the backup structure was necessary, since the deflections were not within specification. The NAA program was now utilized as a design tool and was asked to analyze some 40 HAYSTACK configurations with up to 30 loading conditions each. After five months of analysis, modifications to the backup structure were established that would reduce the deflections to meet the specifications. Computer time totaled approximately 150 hours on an IBM 704, and about 7000 man-hours of support and evaluation were used. Meanwhile, in the same period at Lincoln Laboratory, 25 hours of IBM 7090 time were expended in STAIR runs on 10 to 15 configurations with up to 10 loading conditions per configuration. Approximately 2000 man-hours supported this effort. At the conclusion of this phase of the project, results of both computer programs indicated that it would be possible to predict the

deflections of the reflector within ± 0.005 inch. The maximum deflections due to gravity from face-up to face-side, as computed by NAA and checked by STAIR, was ± 0.040 inch, which was a reasonable value for gravity effects.

To minimize reflector distortion due to thermal effects, all parts of the reflector system were made of one material; namely, aluminum. For design studies, the vertical temperature gradient of the air within the radome was estimated to be approximately 10°F, and in addition, the temperature of the air near the side of the radome facing the sun was estimated to be about 10°F above that of the air near the opposite side. The effect of these gradients on the structure was studied using computer programs at NAA and at Lincoln. These studies predicted that surface distortions as large as ± 0.017 inch would occur with a 10°F gradient across the radome diameter. Later experience has shown that the estimate for the temperate gradient could be exceeded on hot sunny days, and therefore, an air moving system has been incorporated into the radome.

In December of 1962, when the actual backup structure had been fabricated and assembled at the NAA plant, a series of static load tests was carried out. Six separate load configurations were applied to the structure, and deflections measured at preselected points throughout the structure. After reducing raw data to eliminate rigid body rotation and translation of the structure, resulting deflections were compared with both STAIR and NAA computed deflections for the same loading conditions.

A third computer program, developed in the intervening time by IBM, was also used to calculate deflections. This program is called "FRAN," an acronym for FRame ANalysis. The HAYSTACK antenna was one of the initial test problems run on the FRAN program. Figure 9 lists the tests and comparable computed deflections for the X-direction (parallel to the reflector focal axis) loads as analyzed by both Lincoln Laboratory (STAIR and FRAN) and NAA. The reference coordinate system employed is shown in Fig. 10.

Considering that the accuracy of the reduced test deflection data was about ± 0.008 inch, a study of the tables reveals good correlation between measured and computer-predicted results. Ninety-four percent of all the measured test deflections were no more than 0.010 inch away from the MIT-STAIR computed values, and 69 percent were no more than 0.005 inch from computed values.

From the foregoing results, it may be concluded that, for the backup structure alone, the STAIR, FRAN, and NAA programs yield generally good results, and that computer techniques can reliably predict backup structure deflection behavior.

Over the past year and a half, the availability of the IBM FRAN program has allowed the inclusion of the effect of rigid joints in the backup structure (instead of pin-ended joints) and the development of a satisfactory grid analogy. Prior to the release of this program by IBM, the HAYSTACK structure was used to field test the program to evaluate its usefulness as a analytical tool. The backup structure includes many rigid joints; however,

BACK-UP STRUCTURE DEFLECTION COMPARISONS
IN-PLANT TESTS – BACK-UP STRUCTURE ONLY
RF DIRECTION DEFLECTIONS (in.)
(For Target and Load Locations See Fig. 10)

LOAD DIRECTION	TEST	LOAD	RESULTS FROM / NODE POINT (Target Row and Number)		H-1	H-7	H-9	H-17 H-49 (avg)	H-33	F-1	F-17
FACE-UP TYPE LOAD	1	2000 lb H-1 and H-33	TEST		−0.265	−0.052	0.042	0.081	−0.234	−0.102	0.050
			COMP.	NAA	−0.223	−0.054	0.040	0.088	−0.223	−0.102	0.053
				STAIR	−0.223	−0.054	0.040	0.088	−0.241	−0.101	0.053
				FRAN	−0.223	−0.051	0.037	0.082	−0.232	−0.095	0.049
	2	2000 lb H-17 and H-49	TEST		0.066	0.039	−0.025	−0.363	0.075	0.049	−0.099
			COMP.	NAA	0.076	0.045	−0.033	−0.343	0.076	0.053	−0.104
				STAIR	0.076	0.046	−0.033	−0.345	0.078	0.053	−0.104
				FRAN	0.070	0.043	−0.029	−0.333	0.071	0.047	−0.097
	3	1000 lb H-9, H-25, H-41 and H-57	TEST		−0.012	−0.033	−0.032	0.009	−0.006	−0.003	0.007
			COMP.	NAA	−0.005	−0.037	−0.037	0.009	−0.005	−0.002	0.002
				STAIR	−0.005	−0.037	−0.036	0.010	−0.004	−0.002	0.003
				FRAN	−0.006	−0.037	−0.035	0.011	−0.005	−0.003	0.003
FACE-SIDE TYPE LOAD	4	2000 lb H-1 and H-33	TEST		−0.040	−0.009	0.000	0.003	0.043	−0.017	0.000
			COMP.	NAA	−0.041	−0.009	0.002	0.000	0.041	−0.017	0.000
				STAIR	−0.041	−0.009	0.003	0.000	0.051	−0.017	0.000
				FRAN	−0.040	−0.009	0.002	0.001	0.049	−0.016	0.001
	5	2000 lb H-17 and H-49	TEST		−0.015	−0.004	0.002	−0.003	0.012	−0.008	0.000
			COMP.	NAA	−0.015	−0.011	0.000	0.000	0.015	−0.010	0.000
				STAIR	−0.015	−0.010	0.001	0.001	0.016	−0.009	0.000
				FRAN	−0.014	−0.010	0.001	0.001	0.015	−0.009	0.001
	6	1000 lb H-9, H-25, H-41 and H-57	TEST		−0.001	−0.004	−0.017	−0.001	0.008	0.001	−0.003
			COMP.	NAA	−0.008	−0.011	−0.019	0.000	0.008	−0.005	0.000
				STAIR	−0.007	−0.010	−0.018	0.000	0.008	−0.004	0.000
				FRAN	−0.007	−0.010	−0.018	0.001	0.007	−0.005	−0.001

FIG. 9. *Comparison between calculated and measured deflections.*

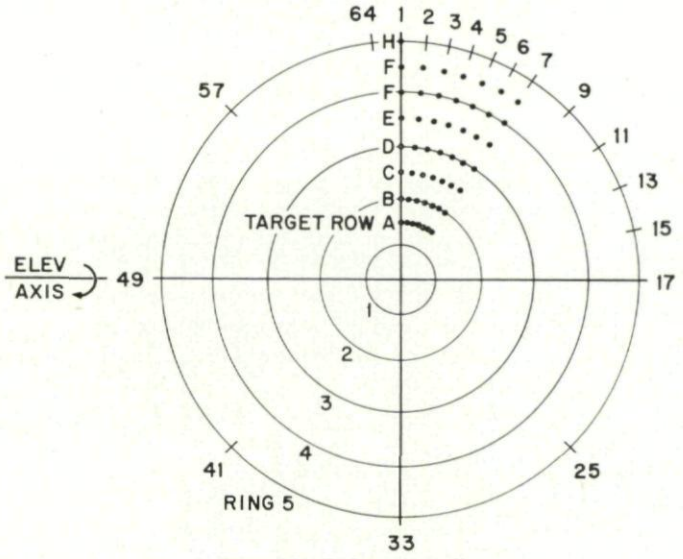


FIG. 10. *Reference coordinate system.*

the significant rigid joints are in ring three and the trunnion beam, where the members are quite stubby (length to diameter ratio ≈ 5) and are butt-welded. The effect of joint rigidity can be seen by comparing the STAIR and FRAN analytical methods. In STAIR, the resistance to displacement is only a function of the cross-sectional area of the member, while FRAN adds to this the additional stiffness resulting from the added constraint that the angular relationship of the members at each joint remain constant. This is a function of the bending and twisting inertia of the member. In addition, FRAN calculates the resulting rotations of the entire joint along with the X-Y-Z displacements. Thus, FRAN solves for twice as many variables at each joint as STAIR, six instead of three. These rotations are not, in general important; however, the added displacement resistance contributed by the

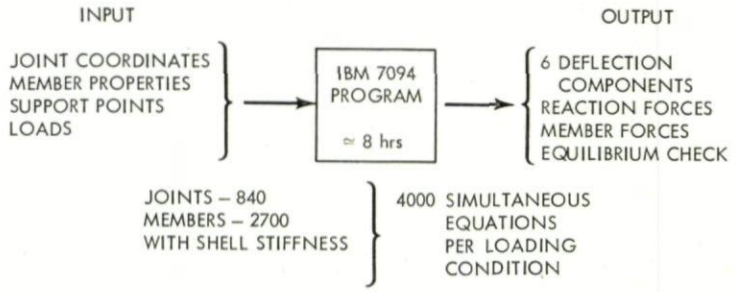


FIG. 11. *FRAN program.*

member bending properties is important, especially since in general it varies throughout the structure. If it were uniform throughout, it would result in reduced over-all deflections. However, it is local in nature, which can cause it to have a significant local effect on surface distortion, even though the over-all deflection change is still small. The basic parameters of FRAN, as programmed for HAYSTACK on the IBM 7094, are indicated in Fig. 11.

One added advantage of the FRAN program is its ability to more adequately predict the behavior of a membrane shell. Early efforts using the model tests for correlation showed that good representations of the membrane and bending stiffness of the shell were not available. Additional theoretical work was done in this area with no really good results until the FRAN program became available. The shell has too many properties to be represented effectively by the single basic parameter available in STAIR, i.e., the cross-sectional area of the member. FRAN's ability to deal with bending makes it possible to develop a more exact analogy.

The FRAN program has been used in three main areas; the in-plant tests without panels, the on-site tests with the complete structure, and the operating condition of gravity loads. The degree of correlation with the various tests, as will be shown, was even better than with STAIR.

The structural assembly, at the time of the in-plant tests during December 1962, included the entire basic ring structure, trunnion beams, and the RF

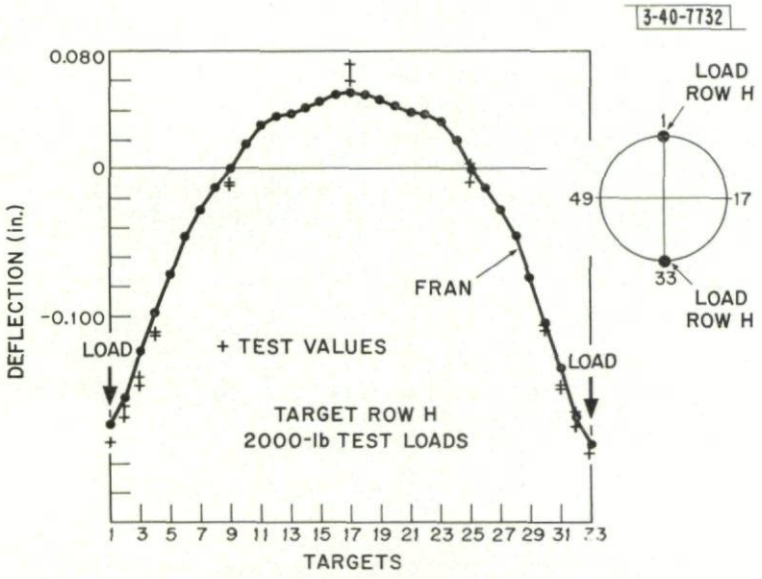


FIG. 12. *Comparison between computed deflection values and on-site tests.*

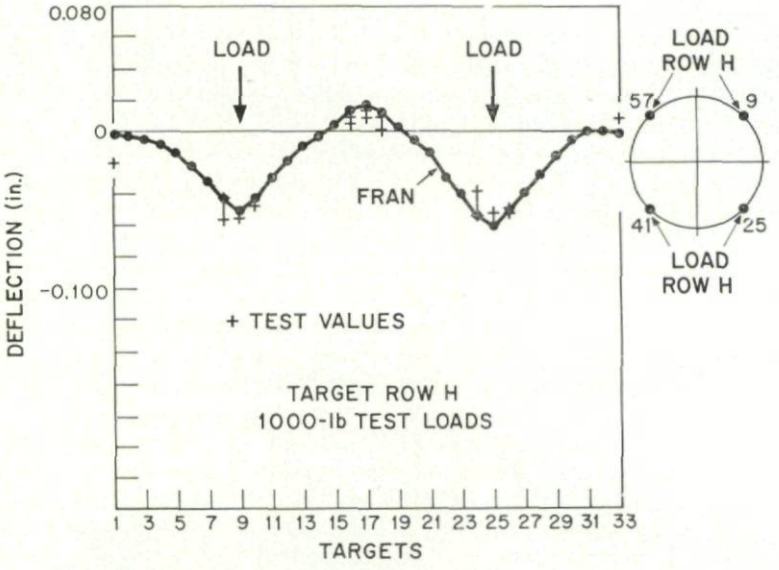


FIG. 13. *Comparison between computed deflection values and on-site tests.*

raceway structure. Portions of the antenna, which were omitted in the assembly, were also left out of the analysis, so that the analytical structure was as close as possible to the actual test structure. The items not included in these tests were the honeycomb surface panels, the quadripod support structure for the secondary reflector, and the tri-ballast counterweight structure along with its associated pendulum ballast beams and cables. The tests consisted of loading a small number of joints on the structure with weights applied through cables. There were six cases with loads applied vertically, as in the face-up position, and six cases with loads applied horizontally. The initial FRAN run for these tests considered the structure to be completely pin-jointed, in order to make a direct comparison with the STAIR analysis. Comparison of these results showed that FRAN gave results identical to the earlier STAIR work within 0.001 inch. The data for this run were then revised to include bending rigidity of the members in ring three and the trunnion beams. The results of this run were then compared with STAIR and the actual tests. The difference between STAIR and FRAN are fairly small, as can be seen in Fig. 9, which is a sample of the comparisons made; the FRAN results, however, are closer than STAIR to the test results for about 70 percent of the measured points. This is extremely good, considering that STAIR was within 0.010 inch of the test values at over 90 percent of the values and 0.010 inch is considered to be the measurement uncertainty.

After disassembly, shipping, and final complete assembly at the site, more comprehensive tests were made. For these tests, the structure was complete and included its integral shell. An analytic representation of the continuous shell by means of a network of beams and bars (a lattice analogy) was developed from the theory of shells and framed structures, and checked using actual shells with known solutions available in various publications. A formulation of a shell attached to a backup structure by means of standoff studs (both the normal and shear studs) through the joints was developed and used to compute deflections. A comparison between computated deflection values based on this more complete analogy and the on-site tests can be seen in Figs. 12 and 13. Most of the analytical values are within 0.010 inch of the test values.

The amount of structural testing on the HAYSTACK system has been much more extensive than for other comparable structures. The results to date have been extremely gratifying and indicate that the analytical tools (such as FRAN and STAIR) now available to structural engineers can provide data which are even more precise than can be obtained by measurement from carefully controlled full-scale tests.

This testing had one major purpose, i.e., to demonstrate the validity of the analytical model for a complete set of tests and to establish a level of confidence for the computer analysis. The test program indicated that the analysis was valid and adequate.

The gravity load cases have been run on the computer and the results indicate general agreement with STAIR in that the peak distortions have remained at ± 0.040 inch. The FRAN predicted pattern of distortion, while

it is essentially the same, does show significant local difference from those obtained from STAIR. The additional effort required by the FRAN computations has proven valuable, since the surface may be adjusted in the face-up orientation to a distorted paraboloidal shape to minimize surface errors when the reflector is tilted to 45°, and errors in the bias rigging table as small as 0.020 inch could increase local distortions by up to 60 percent.

FRAN has been used to evaluate the optimum use of the three variable counterweights that have been added to the structure to permit gravity compensation of sectors of the surface. These are the two pendulum ballast beams which apply cable forces to the top of the antenna as a function of elevation angle, while the third is the outboard counterweight system which modifies the forces on the back-up structure as a function of elevation angle to give a minimum distortion to the surface.

The details of this design effort will be documented in a separate report.

3. RADOME

The desire to locate this new experimental station relatively near the M.I.T. Lincoln Laboratory dictated that the antenna must survive the rigors of a New England environment. Studies indicated that it would not be possible to construct a large, steerable microwave antenna with the required surface tolerance and pointing precision unless it were protected from the wind, snow, and ice. The use of a radome for this purpose appeared preferable to a movable shelter, as it would allow the system to operate a greater percentage of time. To be acceptable, the effective noise temperature of the radome must be low since the new experimental station would employ helium-cooled, low-noise receivers.

Prior to the start of the HAYSTACK program, the large-diameter rigid radomes that were in use employed relatively thick dielectric walls or sandwich-type construction to fulfill the structural requirements. These thick-walled all-plastic radomes exhibited electromagnetic performance which was unacceptable at wavelengths shorter than 30 cm. Large-diameter, air-inflated radomes also suffered from similar electromagnetic limitations at wavelengths shorter than 5 cm and, in addition, did not provide the intrinsic degree of structural reliability that seemed desirable. Studies at the Lincoln Laboratory led to the development of a type of radome employing a spherical structural framework of metal beams which made it possible to use relatively thin, low-loss dielectric panels to cover most of the radome surface area. At the time that the HAYSTACK program was being formulated, a 150-foot-diameter, metal space-frame radome, designed to withstand a 200-mph Arctic environment, was under procurement by the Air Force. This radome, which had the shape of a $\frac{5}{8}$ sphere, was made available to Lincoln Laboratory for electromagnetic evaluation testing. Studies indicated that it would be possible to extend the lower section of the $\frac{5}{8}$ sphere to convert it to a $\frac{9}{10}$ sphere, and that this change could be accomplished without weakening the structure below the wind survival specification of 130 mph that had been established for the Massachusetts location.

An antenna diameter of 120 feet was selected since it was the largest fully steerable parabolic aperture on an elevation-azimuth mount which could be fitted into the radome. With this size, the clearance between the radome and the rim of the reflector is about $3\frac{1}{2}$ feet. Figure 14 shows the placement of this antenna within the radome.

In the original 200-mph radome design, approximately 6 percent of the spherical surface was occupied by the metal space framework. However, since a substantial portion of the electromagnetic energy passes through

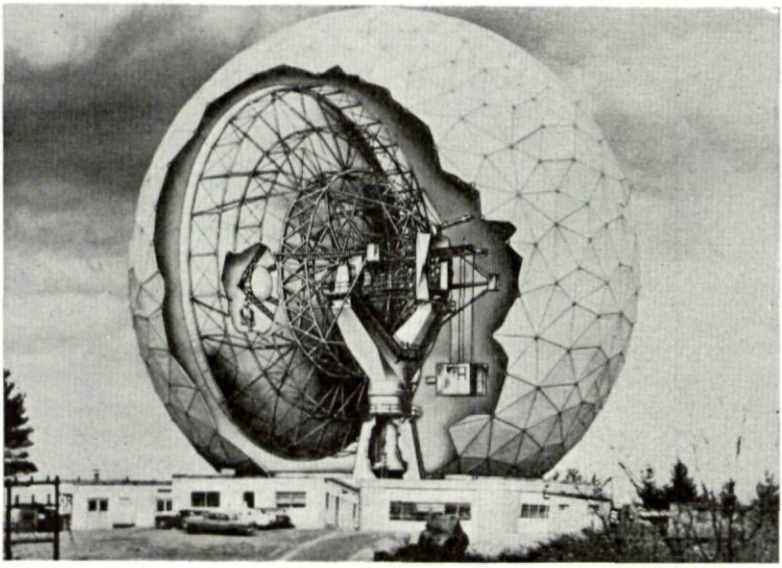


FIG. 14. *Placement of antenna within radome.*

the spherical radome surface at an oblique angle, the shadowing produced by the metal ribs of the radome will be greater than 6 percent. In the HAYSTACK system, the use of a 120-foot-diameter antenna within a 150-foot diameter radome results in a computed effective aperture blockage of about 11 percent. Electromagnetic tests on a model range demonstrated that the noise temperature contribution of the metal space frame was small. The loss of 11 percent of the effective aperture seemed tolerable in order to obtain an improved environment for the remaining 89 percent of the antenna surface.

The metal members of the space framework vary in length between 9 and 15 feet and are aluminum extrusions approximately 3 inches by 5 inches in cross section. This framework is used to support fiberglass dielectric panels 0.032 inch thick. The electrical performance of this radome has been calculated and is shown in Fig. 15. It is noted that throughout the 1- to 10-Gc/s frequency region the loss attributable to the space frame is about 1.1 db. At higher frequencies significant energy is reflected from the membrane, since it has a dielectric constant of about 4. This reflected energy is not due

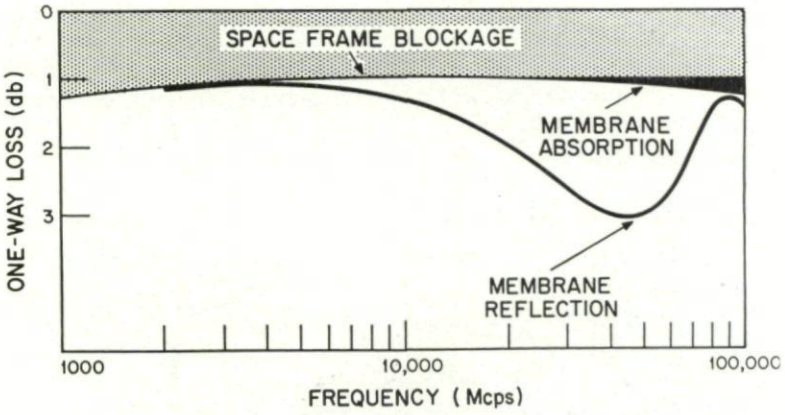


FIG. 15. *Electrical performance of radome.*

to an ohmic loss and therefore does not directly increase the noise temperature of the receiver. At a frequency near 90 Gc/s where the membrane thickness approaches $\lambda/4$, the reflections at the front and rear boundaries of the membrane cancel. The loss tangent of the fiberglass panels is about 0.0025 and this results in an ohmic loss which will contribute to the system noise temperature. This noise contribution for two values of aperture illumination and at two wavelengths is plotted in Fig. 16.

It is clear that the performance of this radome, while perhaps acceptable, is not ideal. Today, it is possible to design a more satisfactory space-frame radome in the 150- to 200-foot diameter class with an over-all loss of perhaps

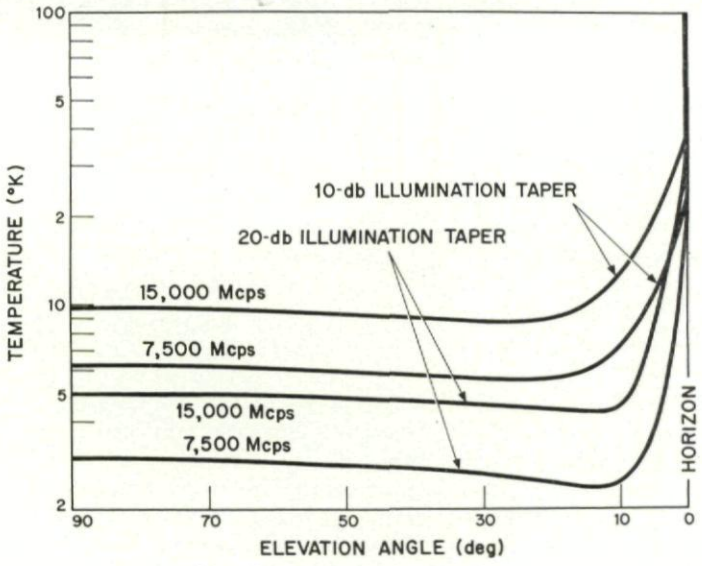


FIG. 16. *Radome noise contribution.*

only 0.5 db. This could be accomplished by designing the structural members for a 130-mph environment, instead of the 200 mph as employed for this particular design, and by constructing the beams from high-tensile-strength steel instead or aluminum. By employing presently available digital computer analytic techniques, a more efficient space-frame design also seems feasible. The use of a radome that is larger in proportion to the antenna diameter would also reduce the electrical shadowing. Radome membrane materials with improved electrical and mechanical properties are now available, and the use of these new materials would reduce the membrane losses by a factor of about 3.

4. PLUG-IN ELECTRONICS EQUIPMENT

In order to facilitate the use of the HAYSTACK system over a wide frequency range and for a variety of experiments, the antenna structure has been designed to accommodate plug-in equipment rooms. These equipment rooms will be used to house low-noise receiver components, cryogenic equipment, and the final amplifier of high-power microwave transmitters. Provision has been made to mount any one of a number of interchangeable equipment rooms directly behind and along the axis of the main parabolic reflector. This equipment configuration is efficient, since it makes it feasible to mount the primary antenna feedhorn directly onto the face of the plug-in room and eliminates the need for rotary joints and long runs of waveguide. An integral hoist system has been provided to raise and lower an 8-foot $\times$ 8 foot $\times$ 12 foot plug-in room with a maximum weight of 7000 pounds. A typical plug-in equipment room is shown in Fig. 17.

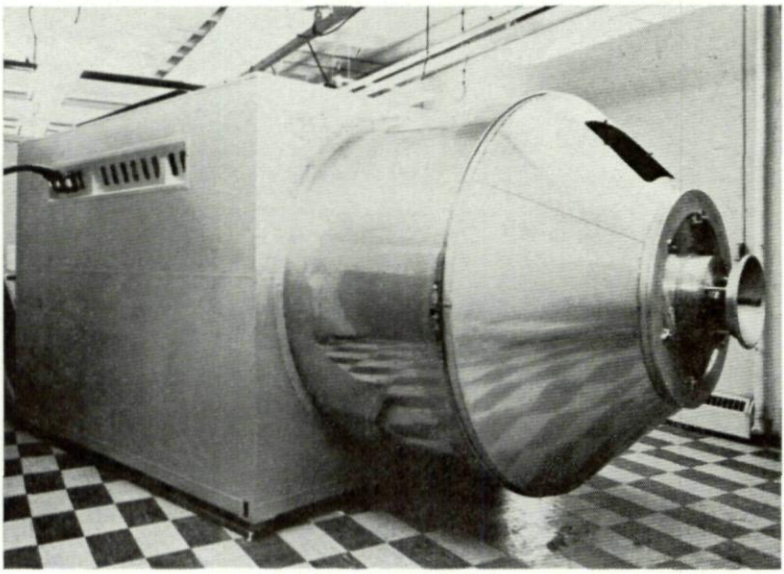


FIG. 17. *Typical plug-in equipment room.*

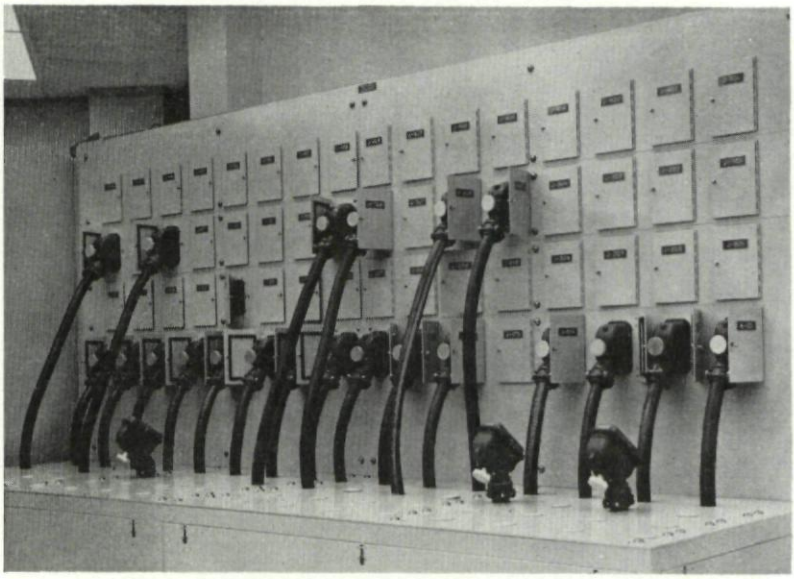


FIG. 18. *Interconnection patch panel.*

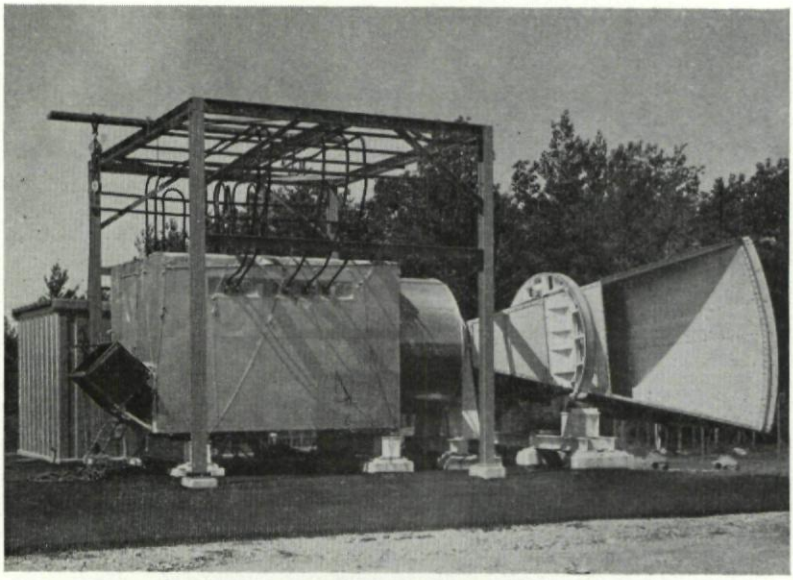


FIG. 19. *Radiometer test stand.*

It is not intended that operating personnel will remain within the plug-in room when the antenna is in use, so provision has been made to control and monitor the electronic equipment remotely. Since the equipment room rotates $\pm 300°$ in azimuth and $90°$ in elevation, cable-wrap systems have been incorporated into the antenna structure. Special cables were procured to obtain the desired compliance and to insure that they would survive repeated flexing.

Provision has also been made to operate the plug-in equipment room at either of two test stands on the ground. A versatile interconnection system has been provided to permit the simultaneous operation of one plug-in room in the antenna while a second plug-in room is under test at the ground stand. High-quality, multiconductor plugs and jacks are employed to allow rapid patching of the approximately 3000 power and control leads and the several hundred coaxial connections that are provided. For transmitter operations, high voltage, compressed air, and up to 260 gallons of cooling water per minute are also available at the plug-in box locations in the antenna and on the test stands. Figure 18 is a photograph of the plug-in patch panel in the main control room. All the connecting plugs are keyed and interlocked.

One of the test stands is situated outside the radome, approximately 150 feet from the main antenna location, to permit radiometric measurement unimpeded by the radome. A photograph of the plug-in equipment room in this test stand is shown in Fig. 19.

5. ANTENNA CONTROL SYSTEM

In order to utilize a high-gain steerable antenna effectively, it is necessary to point the antenna beam with an uncertainty not greater than one-tenth of a beamwidth. In addition, there are experiments which could utilize a beam-pointing capability as precise as $\frac{1}{100}$th of a beamwidth. With the HAYSTACK antenna, where the half-power beamwidth will be $0.05°$ at a wavelength of 3 cm, the fulfillment of even the $\frac{1}{10}$th-beamwidth requirement under dynamic conditions requires a sophisticated control system.

The pointing problem is most severe for low-altitude satellite targets because they pass through the narrow antenna beam at high angular rates. For example, a 200-mile high satellite at closest approach will have an angular rate of approximately $1°$ per second as seen from a ground station, whereas a satellite at a 2000-mile altitude at closest approach will move through the beam at an angular rate of $0.1°$ per second. Targets such as the moon, sun, planets and stars have angular rates of approximately $0.004°$ per second, determined largely by the earth's rotation. Hence, even distant targets of interest, such as radio stars, will pass in and out of the beam in a period of less than 15 seconds. Since the antenna will permit angular rates as high as $3°$ per second, the pointing system must achieve smooth motion with high precision.

In the HAYSTACK system, a high-speed, general-purpose digital computer is used in the (real-time) control system. The computer permits a

relatively unskilled operator to control and direct the antenna in a precise manner. Through the use of a standard typewriter keyboard, the operator can request the antenna to point to designated coordinates or follow one of many pre-planned tracking sequences. Software programs have been prepared which permit an operator to request that the antenna point at such targets as the moon, the planets, and known satellites by simply typing the name of the object in clear text. The computer responds in clear text via the teleprinter, and if necessary, will ask the operator for additional pertinent information. The operator may easily interject modifications to the routine in use and can easily employ one of a number of scan and search modes.

The computer has been programmed to permit the pointing system to compensate for systematic errors in the antenna structure and data-take-off components. At low elevation angles, where the refractive index of the atmosphere will influence the beam pointing position, the computer will be used to introduce corrections to compensate for this effect, and these pointing corrections can readily be modified whenever appropriate. Once an object type has been selected by an operator, a main computer program cycle takes place every 2 seconds. At that time, the new position of the object is determined, superposition of selected scanning takes place, and the position transformed from celestial to radar coordinates. Corrections for refractive and structural deformation are superposed and 4-point interpolation is used to determine the 250 points per second required to provide smooth direction of the antenna servo system.

The computations to find positions of stars, planets, the sun, and the moon are found by preliminary interpolation in stored magnetic tape tabulation of basic *Nautical Almanac* data. Positions of earth satellites and the *West Ford* belt are found by direct evaluation of the satellite equations from a starting point of orbit parameters.

Extra effort has been expended to permit easy use of the facility. As an example, the system can be run in speeded-up time to serve as a planning device; thus, one can easily determine the precise conditions to be expected in azimuth, elevation, range, and Doppler for a particular planetary observation some time in the future. Even while a tracking sequence is in operation, the operator may interrogate the computer memory system to obtain information about other objects whose ephemerides have been stored. The computer also makes it possible for the HAYSTACK system to accept pointing information from remote sources and to make this information useful for directing the HAYSTACK antenna. This involves the generation of parallax and time-delay corrections in real-time. The computer also provides a capability for preparing pointing information to be sent from HAYSTACK to other locations in a form compatible with external sensors.

Computer programs have also been prepared to calibrate the servo-response function of the antenna. A number of interesting new diagnostic techniques are expected to evolve from this flexible capability.

A photograph of the Remington Rand Univac 490 digital computer, as installed at the HAYSTACK facility, is shown in Fig. 20. This computer

was chosen because its word length of 35 binary bits was compatible with the encoder precision and because its computational speeds permitted the generation of pointing commands in real-time from orbit ephemerides. One of the unique features of this computer is its flexible input-output system, which permits the computer to time-share functions. With this capability,

FIG. 20. *Univac* 490 *computer in control room.*

several input and output devices can supply and receive data without interfering with other computer functions.

FREQUENCY CONTROL AND TRANSLATION EQUIPMENT

The frequency translation system is designed to produce precise frequencies for use in exciting various transmitters and for appropriate coherent reference signals for use in the detection and data processing channels.

The basic frequency standards at the station are the Varian model V-4700 rubidium vapor frequency standard and the Hewlett-Packard 107 crystal oscillator. These units are checked against the U.S. national frequency standard by comparison methods, using HF and VLF broadcast time and frequency signals. The long-term accuracy of the frequency standard system as determined by the rubidium unit is 5 parts in 10^{11} per year. The short-term performance of this equipment is 1 part in 10^{11} for a one-second average time. The crystal oscillators are checked against the rubidium standard and serve as flywheels, which can continue in operation in the event of power failure.

Various frequencies are derived as multiples or subharmonics of the standard reference frequencies of 1 and 5 mc/s. These frequencies are used to drive synthesizers and synchronizers which produce frequencies in 10 mc/s increments up to a limit of about 12 Gc/s. For example, the configuration used to provide the stabilized local oscillator for a 7.75 Gc/s radar is shown in Fig. 21. Multiplier chains are used to develop 30- and 50-mc/s harmonics for use as reference standards in the klystron synchronizing system. A 1-watt *X*-band source having essentially monochromatic spectral

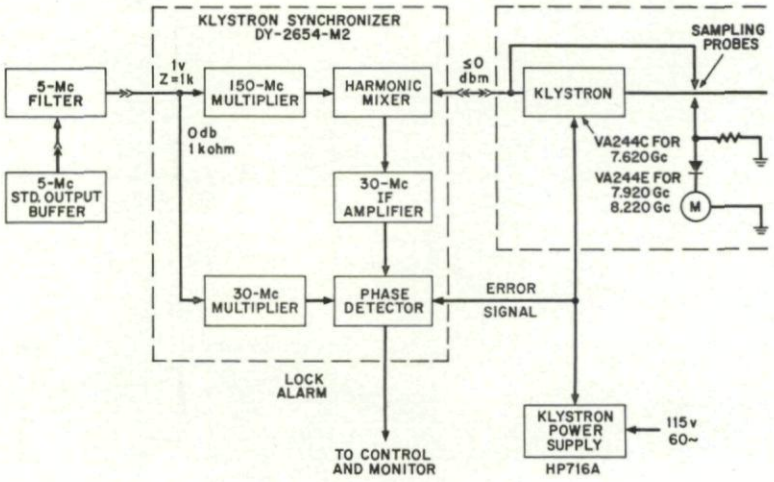


FIG. 21. *Block diagram of local oscillator stabilizing system.*

characteristics with the same intrinsic stability as the crystal oscillator is produced.

Frequency translation equipment shown in Fig. 22 is used to produce the desired transmitter frequency and the IF frequencies required for processing the received signals. The first IF frequency is 130 mc/s. For transmitting, it is synthesized by heterodyning the appropriate harmonics of the standards. Special modulating signals and pulsing can be applied to any of the higher frequency signals. A varactor diode up-converter is used to combine the 130-mc/s IF with the 7.62-Gc/s stabilized local oscillator to obtain the RF excitation for the transmitter. The stalo, of course, serves as the first local oscillator in the receiver chain and heterodynes the received signal down to the 130-mc/s first-IF frequency. A second heterodyne conversion produces an IF of 30 mc/s.

The first-IF amplifier chain, operating with a cooled parametric amplifier in front of it, has a measured dynamic range of 87 db in a 1-kc/s noise band.

A 100-mc/s local oscillator is developed by adding the output of a programmable frequency synthesizer to the 62-mc/s harmonic of the frequency

standard. The Doppler shift of the received signal can be removed using a precomputed ephemeris table with an accuracy of 1 c/s and a precision of 0.1 c/s. Subsequent heterodyne translations are used successively to provide receiver channels with 2-mc/s, 200-kc/s, and 5-kc/s intermediate frequencies,

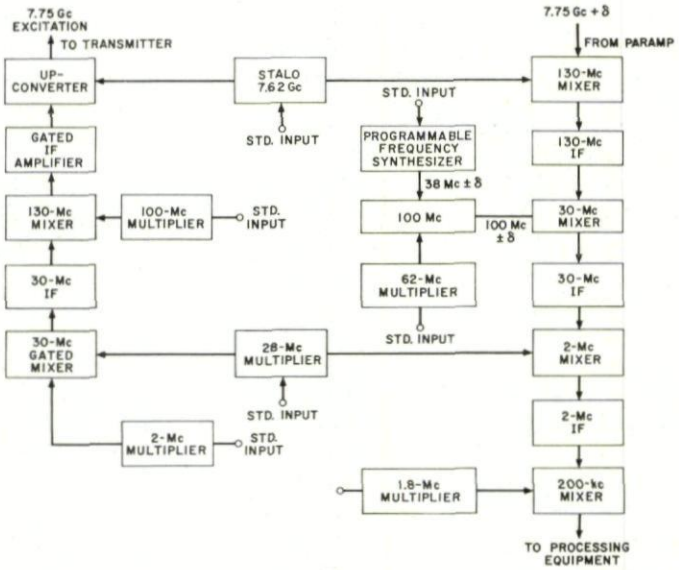


FIG. 22. *Frequency translation system.*

depending on the application. Planetary and lunar radar experiments will use the narrow-bandwidth 5-kc/s intermediate frequency. If the S/N ratio is adequate, a detector system in the 5-kc/s IF channel can be used to develop a phase-locked reference for closed-loop Doppler tracking.

7. THE RADIOMETRIC SYSTEM

The HAYSTACK radiometric system has been designed for research investigations of our own atmosphere, the moon, the planets, stellar radio sources, interstellar gas clouds, and external galaxies. Radiometers operating at frequencies of 5 Gc/s, 8 Gc/s, and 15.5 Gc/s have been completed and radiometers will be added for 35 Gc/s, 1.42 G/cs, (hydrogen line frequency), and 1.67 Gc/s (OH line frequency).

The radiometric system employs several concepts and devices which are new to radio astronomy. A novel approach has been taken in the following areas:

(1) Wide-band tunnel-diode amplifiers have been used for the input amplifier in the 5 Gc/s, 8 Gc/s, and 15.5 Gc/s radiometers. Since radiometer sensitivity is proportional to the system noise temperature divided by the square root of the RF bandwidth, a tunnel-diode

receiver with 1000°K system temperature and 1000-mc/s bandwidth is equivalent to a maser or paramp receiver with 100°K system temperature and 10-mc/s bandwidth. The reliability, stability, and absence of cryogenic equipment makes the use of tunnel diodes very attractive.

(2) As discussed in Section 4, the plug-in antenna room makes it relatively easy to maintain and change radiometer equipment.

(3) A precision square-law detector, synchronous detector, and integrator have been designed for use with high-accuracy digital processing. The synchronous detector and integrators are solid-state and have a drift which is 0.01 percent of maximum output.

(4) A flexible data system has been utilized which can couple any combination of seven 30-bit digital data sources and sixteen analog data sources to a computer, a paper-tape punch, or a printer at a fast, buffered rate. In addition, fifty sets of analog data, representing monitor points in the system, can be fed into any of the above devices at a slow unbuffered rate.

(5) Real-time processing of radiometric data will be accomplished in the digital computer that is also employed for antenna pointing. A program is completed which calibrates the radiometer output and prints and plots antenna temperatures. The computer also checks the fifty radiometer monitor points to see if they fall within limits. Extensive use of the on-line computer for radiometer control, monitoring, and data analysis is expected.

(6) The spectral-line receivers will utilize an accurate and versatile digital autocorrelator for spectral analysis. A 100-channel, 10-mc/s clock-rate correlator is under construction.

A photograph of the radiometer equipment room is shown in Fig. 17. The circular extension on the front of the room supports the antenna feed. Future cryogenic front-end equipment will be mounted in this circular extension, in close proximity to the antenna feed.

Normally, the complete RF portion of a radiometer will be installed in the equipment room and the synchronous detectors, integrators, and data processing equipment will be located in the control room. However, a complete radiometer system with analog pen recorder output can be operated from the equipment room for testing purposes.

The temperature environment of radiometric front ends is extremely critical. For example, if a front-end component has a small loss of 0.25 db and its temperature changes by 2°C, a 0.1°K spurious signal will be introduced into the radiometer. Since the inherent sensitivity of the radiometers in use is of the order of 0.01°K for $\frac{1}{2}$ hour of integration, this is a serious error. A high degree of temperature stability has been achieved in the radiometer room by using proportionally controlled heaters in conjunction with an air conditioner. The heater control loop senses the temperature of the air entering the room and holds this to within ± 0.1°C for one-hour periods and ± 0.25°C for 24-hour periods. The room is well insulated and good radiometer stability has been obtained.

A photograph of a typical radiometer RF components control panel is shown in Fig. 23. In normal operation the feed switch is in the antenna position, the comparison noise source and calibration noise source are off, the balance noise source is on, the ferrite switch is modulated at a 40-c/s rate, and the RF attenuator is in the normal (no attenuation) position. The ferrite switch is then switching the receiver input between the comparison load (a thermally insulated termination at 300°K) and the antenna with

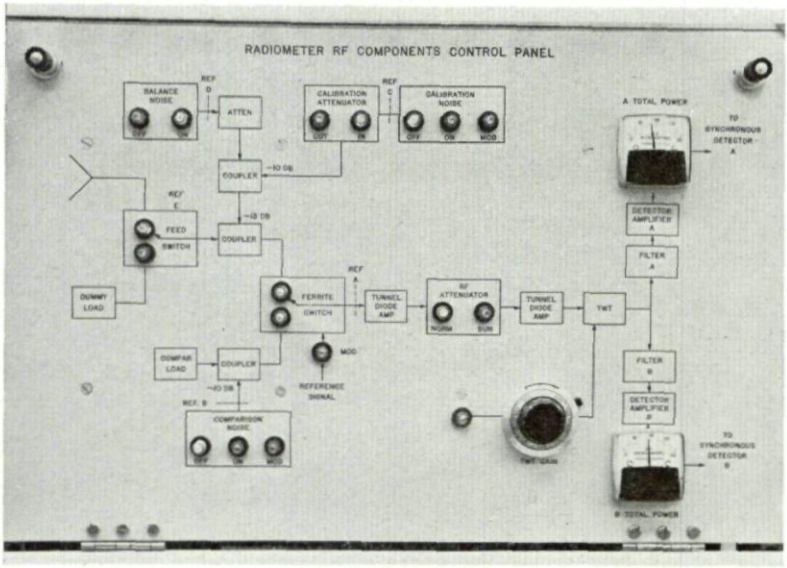


FIG. 23. *Radiometer control panel.*

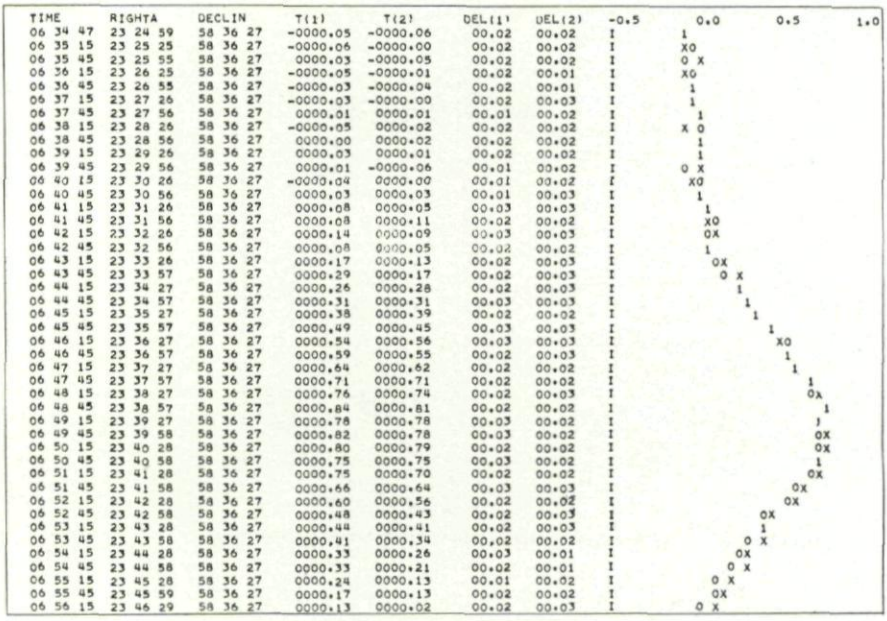
balance noise (−300°K) added. A 40-c/s signal whose amplitude is proportional to antenna temperature is thus produced at the detector output, where it is amplified, synchronously detected, and integrated.

The basic receiver is of the TRF type. A 4-stage tunnel-diode amplifier followed by a single traveling-wave tube amplifier is used. The TWT stage is needed because the maximum linear output of present tunnel-diode amplifiers is not quite enough to properly drive the diode detectors. The characteristics of the amplifier chain are: 60 db gain, 1000 mc/s bandwidth, and system noise temperatures of 1000°K at 5 and 8 Gc/s and 2500°K at 15 Gc/s. The output of the amplifier chain is divided into two 500-mc/s bandwidth bands (for interference detection purposes) before detection.

The link between the radiometers and the data recording system is through a synchronous detector-integrator designed for operation on radiometer signals which are to be digitally processed. The design goal of this unit is drift and noise levels which are 0.01 percent of maximum output because the post-integration signal-to-noise ratio on strong sources will approach

10,000 to 1 and the A/D conversion equipment has an accuracy of 0.01 percent. In addition, the wide dynamic range synchronous detector makes scale changes unnecessary and simplifies observations.

The synchronous detector contains both an RC integrator for analog pen recorder display and a true finite-time integrator for digital recording equipment. The true integrator is sampled by a high-speed A/D converter at the conclusion of the integration period which can be varied between 0.3



TIME	RIGHTA	DECLIN	T(1)	T(2)	DEL(1)	DEL(2)	-0.5	0.0	0.5	1.0
06 34 47	23 24 59	58 36 27	-0000.05	-0000.06	00.02	00.02	I	1		
06 35 15	23 25 25	58 36 27	-0000.06	-0000.00	00.02	00.02	I	XO		
06 35 45	23 25 55	58 36 27	0000.03	-0000.05	00.02	00.02	I	0 X		
06 36 15	23 26 25	58 36 27	-0000.05	-0000.01	00.02	00.01	I	XO		
06 36 45	23 26 55	58 36 27	-0000.03	-0000.04	00.02	00.01	I	1		
06 37 15	23 27 26	58 36 27	-0000.03	-0000.00	00.02	00.03	I	1		
06 37 45	23 27 56	58 36 27	0000.01	0000.01	00.01	00.02	I	1		
06 38 15	23 28 26	58 36 27	-0000.05	0000.02	00.02	00.02	I	X 0		
06 38 45	23 28 56	58 36 27	0000.00	0000.02	00.02	00.02	I	1		
06 39 15	23 29 26	58 36 27	0000.03	0000.01	00.02	00.02	I	1		
06 39 45	23 29 56	58 36 27	0000.01	-0000.06	00.01	00.02	I	0 X		
06 40 15	23 30 26	58 36 27	-0000.04	0000.00	00.01	00.02	I	XO		
06 40 45	23 30 56	58 36 27	0000.03	0000.03	00.01	00.03	I	1		
06 41 15	23 31 26	58 36 27	0000.08	0000.05	00.03	00.03	I	1		
06 41 45	23 31 56	58 36 27	0000.08	0000.11	00.02	00.02	I	XO		
06 42 15	23 32 26	58 36 27	0000.14	0000.09	00.03	00.03	I	0X		
06 42 45	23 32 56	58 36 27	0000.08	0000.05	00.02	00.02	I	1		
06 43 15	23 33 26	58 36 27	0000.17	0000.13	00.02	00.03	I	0X		
06 43 45	23 33 57	58 36 27	0000.29	0000.17	00.02	00.03	I	0 X		
06 44 15	23 34 27	58 36 27	0000.26	0000.28	00.02	00.03	I	1		
06 44 45	23 34 57	58 36 27	0000.31	0000.31	00.03	00.03	I	1		
06 45 15	23 35 27	58 36 27	0000.38	0000.39	00.02	00.02	I	1		
06 45 45	23 35 57	58 36 27	0000.49	0000.45	00.03	00.03	I	1		
06 46 15	23 36 27	58 36 27	0000.54	0000.56	00.03	00.03	I	XO		
06 46 45	23 36 57	58 36 27	0000.59	0000.55	00.02	00.03	I	1		
06 47 15	23 37 27	58 36 27	0000.64	0000.62	00.02	00.02	I	1		
06 47 45	23 37 57	58 36 27	0000.71	0000.71	00.02	00.02	I	1		
06 48 15	23 38 27	58 36 27	0000.76	0000.74	00.02	00.03	I	0X		
06 48 45	23 38 57	58 36 27	0000.84	0000.81	00.02	00.02	I	1		
06 49 15	23 39 27	58 36 27	0000.78	0000.78	00.03	00.02	I	1		
06 49 45	23 39 58	58 36 27	0000.82	0000.78	00.03	00.02	I	0X		
06 50 15	23 40 28	58 36 27	0000.80	0000.79	00.02	00.02	I	0X		
06 50 45	23 40 58	58 36 27	0000.75	0000.75	00.03	00.02	I	1		
06 51 15	23 41 28	58 36 27	0000.75	0000.70	00.02	00.02	I	0X		
06 51 45	23 41 58	58 36 27	0000.66	0000.64	00.03	00.03	I	0X		
06 52 15	23 42 28	58 36 27	0000.60	0000.56	00.02	00.02	I	0X		
06 52 45	23 42 58	58 36 27	0000.48	0000.43	00.02	00.03	I	0X		
06 53 15	23 43 28	58 36 27	0000.44	0000.41	00.02	00.03	I	1		
06 53 45	23 43 58	58 36 27	0000.41	0000.34	00.02	00.02	I	0 X		
06 54 15	23 44 28	58 36 27	0000.33	0000.26	00.03	00.01	I	0X		
06 54 45	23 44 58	58 36 27	0000.33	0000.21	00.02	00.01	I	0 X		
06 55 15	23 45 28	58 36 27	0000.24	0000.13	00.01	00.02	I	0 X		
06 55 45	23 45 59	58 36 27	0000.17	0000.13	00.02	00.02	I	0X		
06 56 15	23 46 29	58 36 27	0000.13	0000.02	00.02	00.03	I	0 X		

FIG. 24. *Sample of high-speed printer output.*

and 30 s. The high-speed A/D converter can be used to sample many true integrators. The synchronous detector contains a post-detection gain modulator which is an innovation to radiometry. The synchronous detector is all solid-state, utilizing approximately 75 transistors.

Radiometer data are collected in two types of scans: a *data scan* for the main radiometer data and an *auxiliary scan* for monitoring data. In the *data scan*, 16 sources of analog data are multiplexed into an A/D converter and the analog voltages are converted into 17-bit digital numbers. In addition, 7 sources of digital information are scanned. These 23 words of data are collected in 23 ms and are stored in a circulating delay-line memory. Any combination of these 23 data words can then be fed into either a printer, a papertape punch, or the Univac 490 computer. A new *data scan* is initiated periodically at adjustable intervals between 0.1 s and 100 s. The *auxiliary scan* consists of a stepping-switch which will scan up to 50 sources of analog

monitor data. This information is converted in the same 17-bit A/D converter as the analog *data scan* data and can be fed into either a printer, a paper-tape punch, or the computer.

An initial radiometer data processing program has been written which provides calibrated values of antenna temperatures in real time. A sample of a high-speed printer output from this program is shown in Fig. 24. From left to right the columns of printer output represent time, right ascension, declination, antenna temperatures in data channels 1 and 2, and rms fluctuations of these antenna temperatures. The right-hand side of the printer page is devoted to a plot of antenna temperatures.

8. TRANSMITTER POWER SUPPLY AND HEAT EXCHANGER

The HAYSTACK plug-in equipment room concept has made it feasible to operate over a wide frequency spectrum in both radar and communication modes. For this reason, it was important to provide a power supply and cooling system which could operate with a variety of transmitter tubes, including klystrons, traveling-wave tubes, and amplitrons, in CW and pulsed applications. In an attempt to obtain reasonable balance between the cost of the transmitter and the other relatively expensive components necessary in the experimental system, it appeared appropriate to install a versatile high-voltage power supply with approximately a 1 megawatt DC output capability. This was greater than needed to supply the microwave transmitter power tubes that were available in 1960, but seemed approximately correct for the improved transmitter tubes expected during the next decade. For this reason, a power supply and cooling system have been specified which are somewhat larger and more versatile than those commonly found in typical radar and communication systems.

The specifications of the power supply may be summarized as follows:

(1) Average Power: The high-voltage DC supply shall be capable of producing, on a continuous basis, 1000kW DC at any voltage from 20 kV to 120 kV.

(2) Ripple: The maximum peak-to-peak ripple of the power supply DC output voltage shall be $\frac{1}{2}\%$ of the output voltage under any combination of voltage and power output.

(3) Energy Storage: Sufficient energy storage shall be included to limit the voltage drop to 4% of the initial voltage when the supply is delivering a 2-ms. pulse containing 10,000 joules of energy.

(4) Pulsewidths: The supply shall be capable of delivering energy to a modulator in the form of pulses of up to 10,000 joules at pulsewidths of 2 ms. and greater.

(5) Pulse Repetition Rates: The supply shall be capable of delivering its energy to a modulator at any pulse repetition rate at which the average rating of the supply is not exceeded.

(6) Voltage Changes: The output voltage shall be capable of continuous adjustment from zero to any voltage in the specified operating range.

Power supply components may be connected in various configurations to allow the attainment of the full output power within each of several voltage ranges.

(7) Polarity: The power supply shall be capable of operating with either the negative or positive output terminal grounded.

The power supply was procured by the Air Force from Energy Systems, Inc., Palo Alto, California (formerly Radiation at Stanford). In the ESI design, the required flexibility is achieved by the use of a plate transformer

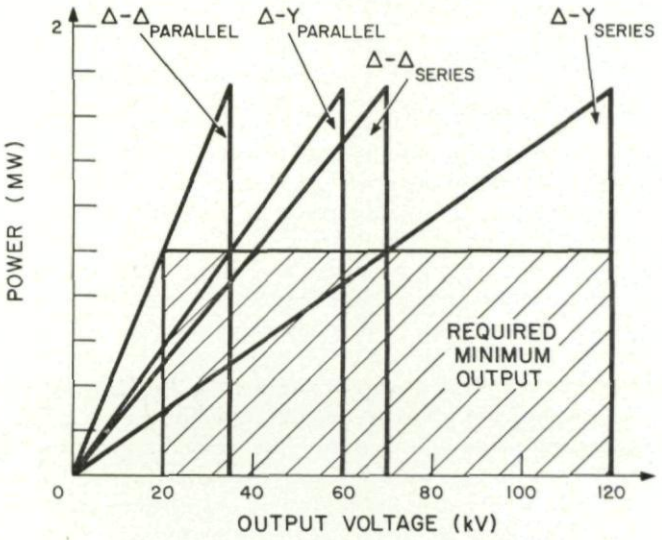


FIG. 25. *Capability of power supply as a function of voltage.*

which has three primary windings and six independent secondary windings. Four of the different transformer connection configurations which are available will be used. There are $\Delta - \Delta$ and $\Delta - y$, with the secondary winding connected in parallel for low-voltage, high-current operation and in series for high-voltage, low-current operation. The four ranges of voltages thus available are 0–35 kV, 0–60 kV, 0–70 kV, and 0–120 kV. Since the supply must be capable of delivering 1000 kW at 20 kV, there is actually more than 1000 kW available at various other voltage levels, as indicated in Fig. 25.

The rectifiers used in this supply are semiconductor diode stacks. The diodes were chosen because they could supply the required load current with large overcurrent capability. This overcurrent capability is an important consideration for the prevention of damage to a rectifier string in the event of power supply faults or external circuit failures. The peak current in a rectifier stack is limited by the reactance of the transformer windings. The reactance has been chosen both to limit the current and to serve as part of a filter to meet the ripple requirements without the use of a separate choke. The use of semiconductor rectifiers eliminates the need for high-reactance,

heater isolation transformers which would be required if thermionic diodes were used. In addition, the interconnection of the rectifiers for operation at the various voltage levels and power supply polarity is simplified.

The 4160 V, 3 phase, 60 cycle input to the power supply goes through the usual fused disconnect switch and high-speed vacuum switches, which should respond in less than one cycle. An Inductrol is used to control the input to the plate transformer, allowing essentially continuous voltage variation. A

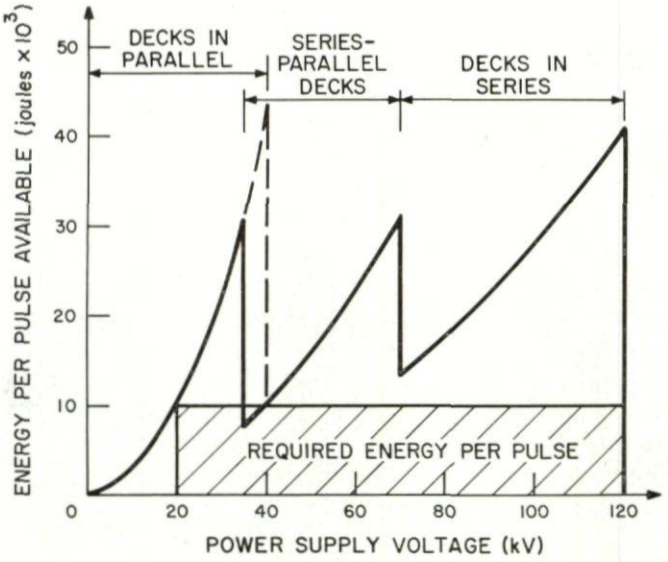


FIG. 26. *Energy per pulse available for 4-percent voltage drop.*

DC feedback regulator loop on the Inductrol serves to regulate the voltage output of the supply to within $\pm 1\%$ over a 25 to 100% output range.

The energy storage bank is composed of three decks, each consisting of 168 capacitors rated at 1.25 μf each at 40kV. These decks may be connected in parallel, series-parallel, or series configuration. By so doing, one may obtain 630 μf at 40 kV with all decks in parallel, 70 μf at 120 kV with all decks in series, and 157 μf with a series-parallel combination of decks. The energy storage per pulse available within the constraint of 4% voltage drop at the various voltage ranges of the power supply is depicted in Fig. 26. The capacitor bank is designed with mechanical shorting bars which discharge the capacitors whenever the system is de-energized. The individual capacitors have series current-limiting resistors to prevent catastrophic discharging of a deck of the bank into a faulty capacitor on that deck, and also to limit the peak current which can be drawn under fault conditions. Figure 27 shows the capacitor bank.

A test load is provided, so that power-supply operation can be verified over its entire range. The load consists of several water columns in parallel. The

resistivity of the column is controlled by varying the salinity of the solution flowing in the columns. There is a heat exchanger associated with the dummy load which transfers the heat generated in the dummy load into the main system heat exchanger.

The system heat exchanger is a double closed-loop system. A distilled water loop is provided for cooling the transmitter tubes and components. An intermediate liquid-to-liquid exchanger transfers the heat from the closed-loop distilled-water system to an ethylene-glycol system which includes a

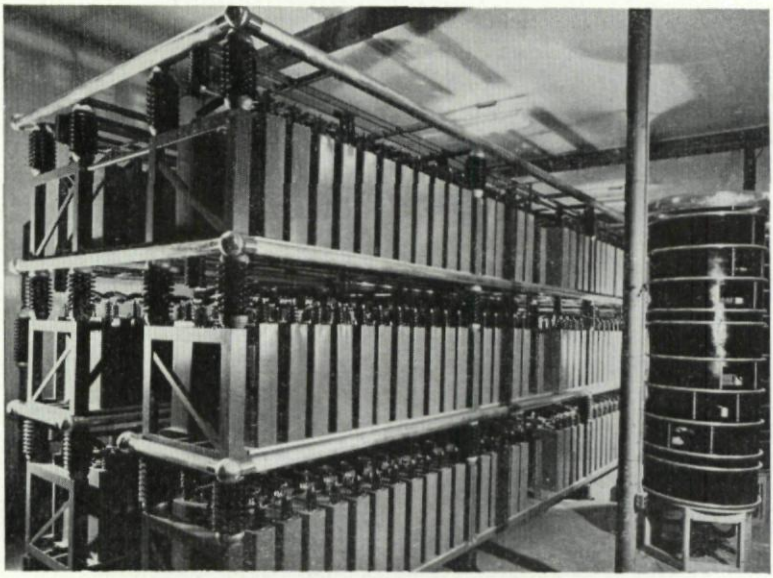


FIG. 27. *Capacitor energy storage bank.*

liquid-to-air heat exchanger. The heat exchange problem in this system is not a simple one, since the major heat source, the transmitter, is located on the antenna a considerable distance away from the intercooler. A highly effective temperature control system is needed to minimize phase fluctuations of the transmitted signal due to variations in the body-cooling water temperature.

In a system with 500,000 joules of stored energy, a fast-acting protective circuit is essential to remove the DC from the transmitter in the event of an arc. A crowbar circuit is used which is proprietary to Energy Systems, Inc. It is a two-ball gap which has a triggering "needle" located on the equipotential plane between the two gaps. The input-logic circuits are designed to operate in the event of excess tube current (an arc), capacitor failure, or capacitor bank overvoltage. One of the important features of the crowbar triggering system is that it provides a number of sequential triggers closely spaced in time. This prevents damage to the system which could result were the crowbar arc to extinguish before the primary power circuits opened.

9. THE INITIAL HIGH-POWER COMMUNICATIONS-RADAR SYSTEM

The initial high-power *X*-band plug-in equipment room has been designed for Venus radar experiments and for communications experiments requiring simultaneous transmission and reception. This plug-in room contains:

(1) the microwave portions of the frequency synthesizer
(2) a 100 kW, 7750-mc/s transmitter,
(3) high power microwave circuits and feed system,
(4) low noise receivers.

A simplified block diagram is shown in Fig. 28.

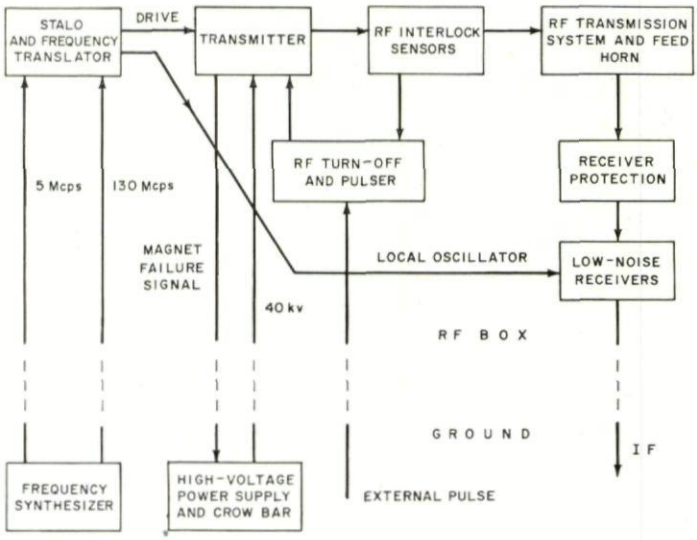


FIG. 28. *Simplified diagram of transmit-receive system.*

The microwave portion of the frequency synthesizer consists of a frequency multiplier and phase-locked klystron and a single side-band modulator which provides a drive signal and local oscillator signal offset by the 130 mc/s intermediate frequency. Local oscillator signals at 7620 mc/s and transmitter drive signals at 7750 mc/s are derived from a 5 mc/s crystal oscillator with appropriate multiplication. A reflex klystron, phase-locked to the multiplied 5 mc/s signal, supplies the local oscillator signal directly. The drive signal is developed by a frequency translator which offsets the drive from the local oscillator by the 130 mc/s IF frequency by the use of microwave single side-band modulator with appropriate filters.

A VA-617 traveling wave tube is used to drive a VA-879 5-cavity klyston final amplifier (Fig. 29). This klystron, recently developed by Varian Associates, Palo Alto, California, has an output power of 100 kW CW, a

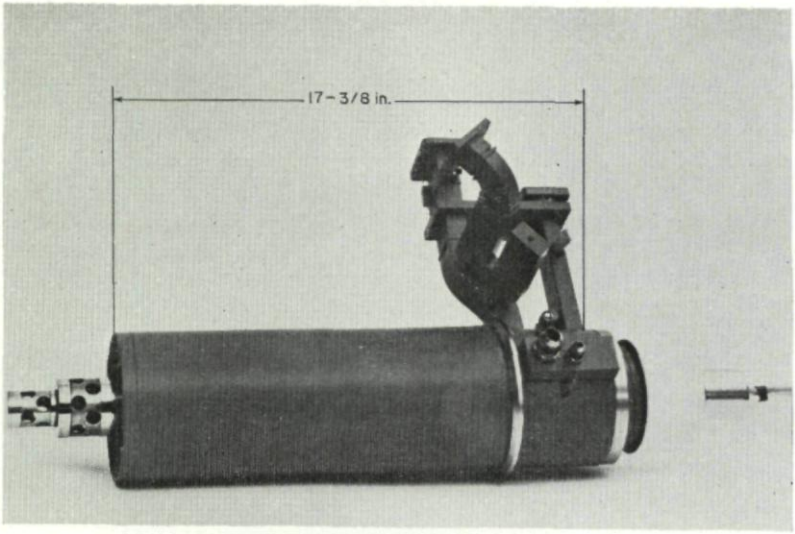


FIG. 29. VA-879 100-kW *klystron.*

gain greater than 50 db, a 3 db bandwidth of about 40 mc/s, and an over-all efficiency greater than 40 percent.

Figure 30 shows the principal features of the RF amplifier chain. Of special interest are the protective circuits which have been incorporated to prevent damage to the tube if an arc should occur in the waveguide system. Two types of protection are provided. (a) against an abnormally high

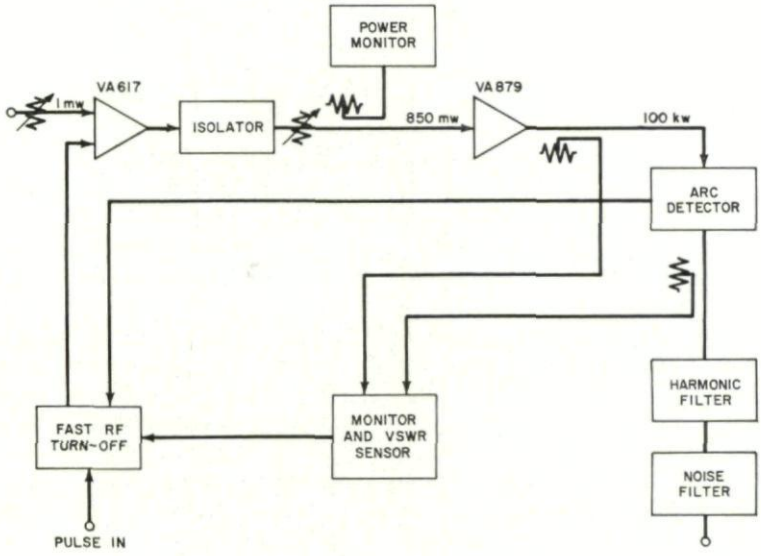


FIG. 30. *Transmitter amplifier chain.*

VSWR, and (b) against an arc which might travel toward the window. A high VSWR might result in internal damage to the tube, and an arc arriving at the window would almost certainly result in window damage which would destroy the vacuum. A microwave circuit is used for the VSWR interlocks; the arc detector uses an optical interlock circuit which would detect an arc in the vicinity of the window. Either of these two circuits will feed a pulse to the fast-turn-off circuits associated with the VA-617 TWT driver. The drive signal supplied to the VA-879 can be reduced 45 db in less than 3 μs and 80 db in less than 5 μs. This same circuit can also be used for pulsing the drive.

All waveguide in the HAYSTACK RF system is of oxygen-free high-conductivity copper, with a few exceptions where available components were of aluminum. All high-power lines are of WR-137 size waveguide up to the power splitter which feeds the monopulse horn. This large waveguide was chosen in order to provide a greater margin or safety against voltage breakdown without going to a nonstandard size of transmission line. In the high-power waveguide all flanges are the CPRF type, made from stainless steel, and are used with a copper gasket which has been designed to provide the best possible contact between mated pieces of waveguide. In resonant-ring tests using this type of flange joint, power levels of over 500 kW CW have been transmitted in WR-137 waveguide without breakdown.

In spite of the low loss of WR-137 OFHC copper waveguide (approximately 0.015 db per foot), the ohmic loss in the walls is sufficient to raise the waveguide temperature to intolerable values when 100 kW CW is transmitted. In order to overcome this difficulty, a copper tubing is attached to opposite waveguide walls by means of solder or heat-conducting epoxy mixtures (usually epoxy containing a large amount of powdered metal which has high thermal conductivity) and water is passed through the tubing. This technique reduces the waveguide surface temperature to very nearly the temperature of the cooling fluid.

In the communication mode of operation, where transmission and reception occur simultaneously, there was concern that the second harmonic of the transmitted signal leaking to the receivers might be of sufficient amplitude to impair reception. With the parametric amplifiers being used, the idler frequency is close to the second harmonic of the transmitted signal so that any incoming second harmonic could be supported by the idler circuits and cause possible saturation effects or undesirable modulation products.

A General Electric Company leaky wall filter was procured for use in this mode of operation and was designed to remove the second and all higher harmonics. This filter is water-cooled, as are all the high-power components, and has an insertion loss at the transmitter frequency of 0.1 db. It provides approximately 50 db attenuation at the second harmonic. This unit has been high-power tested and broke down at 200 kW CW.

In order to reduce the interference produced by transmitter-generated noise at the receiving frequency, a high-power noise-rejection filter was designed and a prototype has been built and successfully tested. The rejection band, tuned to the receiver frequency, provides 60 db rejection over

8 mc/s band, while the passband at the receiver frequency has less than 0.03 db loss. Voltage breakdown occurred at 175 kW CW in resonant-ring tests.

The design and construction of a 100-kW CW circulator which would operate over the band of 7700 to 8400 mc/s was contracted to the Raytheon Company. Their product was in essence two circulators in parallel. A circulator has been successfully high-power tested to 115 kW CW in a resonant ring. The insertion loss is 0.2 db at 7750 mc/s after temperature stabilization. The cold loss is about 0.28 db.

Most of the low-level microwave circuits are in WR-112 waveguide because of the availability of components at the frequency band of interest. These circuits include the monopulse comparators, all of the low-level drive circuits, and the circuits at the input of the receivers.

Two types of receiver protection are required: (*a*) for radar experiments, when the transmitter will be turned off while receiving, and (b) for communication experiments, when simultaneous transmission and reception will be required. For Venus radar experiments, the times involved are so long that mechanical switches will be used for receiver protection. Waveguide switches with electrically driven actuators have been procured and tested. These provide a minimum of 80 db isolation between the input and the disconnected terminal. It is expected that the circulator and the orthogonal mode transducers will both provide a minimum of 20 db additional isolation so the maximum peak power incident on the receivers should be 10^{-5} watts.

For more conventional radar applications, where the transmitted pulse might be 2 to 5 ms long and the range a few hundred or a few thousand miles, the mechanical switches are much too slow. A gaseous discharge attenuator has been designed which will provide 80 to 100 db of protection. The low-level insertion loss has been measured to be less than 0.05 db when cold. The recovery time is about 300 μs to the 3 db point and about 450 μs to complete recovery.

For the communications mode, diplexers have been designed for the purpose of separating the 7750 and 8350 mc/s bands. For early experiments, only the 8350 frequency will be used on reception, so the function of the diplexers is to provide 30 db of transmitter signal rejection in the 8350 mc/s circuits.

Reflection filters for the purpose of reducing the amount of transmitter power leaking to the receiver were procured from Rantec, Inc. The rejection properties of these filters were improved by retuning, so that over 90 db of rejection has been achieved over the transmitter bandwidth. With these improved filters it may be possible to eliminate the diplexers for experiments that do not require reception at the transmitter frequency.

The HAYSTACK antenna employs the Cassegrain type of microwave optics, fed by a horn located near the vertex of the 120-foot-diameter parabolic reflector. A feed horn has been designed to provide inputs to a monopulse comparator. The system is designed to transmit right-hand circular polarization and receive both senses of circular polarization. Monopulse tracking will be done only with the left-hand circularly polarized received signal.

Figure 31 is a block diagram of monopulse system. The transmitter power is fed into a four-way power divider, into the orthogonal mode transducers (OMT), through the circular polarizers, and into the multimode horns shown in Fig. 32.

The multimode horn, fed by four square waveguides, consists of several sections: (*a*) a transition from four waveguides to a large square cross section (this is perhaps the most critical element in the horn structure); (b) a length of straight waveguide of the large square cross section; and (c) a pyramidal horn.

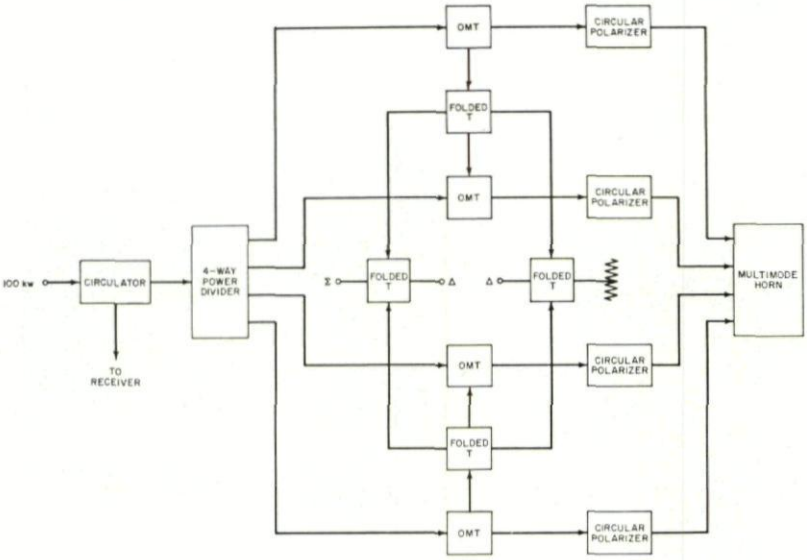


FIG. 31. *RF monopulse system.*

The transition from the four square waveguides to the large square cross section is designed to excite the modes necessary for obtaining angle information and suppressing undesirable modes which may upset the antenna pattern. The transition designed for use in the HAYSTACK antenna achieves this goal. Although not all undesirable modes are completely suppressed, the feed has good patterns in the band 7700 to 8400 mc/s.

The TE_{30} mode is generated by the presence of the short circuit in the aperture plane of the four square waveguides. The configuration of the transition to square waveguide was designed to generate an equal amount of TE_{30} mode out of phase with that generated at the four waveguide aperture. At the same time, it was necessary to avoid generation of undesirable higher-order TE and TM modes. The fin-loading of the transition accomplishes this.

The feed is well-matched at the transmitter frequency, 7750 mc/s, and has a VSWR of 1.2 at 8350 mc/s corresponding to a reflection loss of less than 0.05 db. Compared to a four-horn monopulse feed, the forward gain of this

feed is greater by nearly 1 db over the band of interest. The side lobes of the sum pattern are over 18 db below the main lobe in both planes, compared to 7.5 db and 12.5 db for a four-horn monopulse feed.

The purpose of the straight square waveguide section is to adjust the relative phase of the various modes at the horn aperture. This can be done reasonably well for the principal modes, TE_{10}, TE_{20}, (TE_{11}, TM_{11}) degenerate pair and the (TE_{21}, TM_{21}) degenerate pair. The TE_{01} mode represents the forward or sum signal, and the TE_{20} and (TE_{21}, TM_{21}) degenerate pair form the difference modes. The (TE_{21}, TM_{21}) degenerate provides the E-plane taper which makes possible the reduction of E-plane

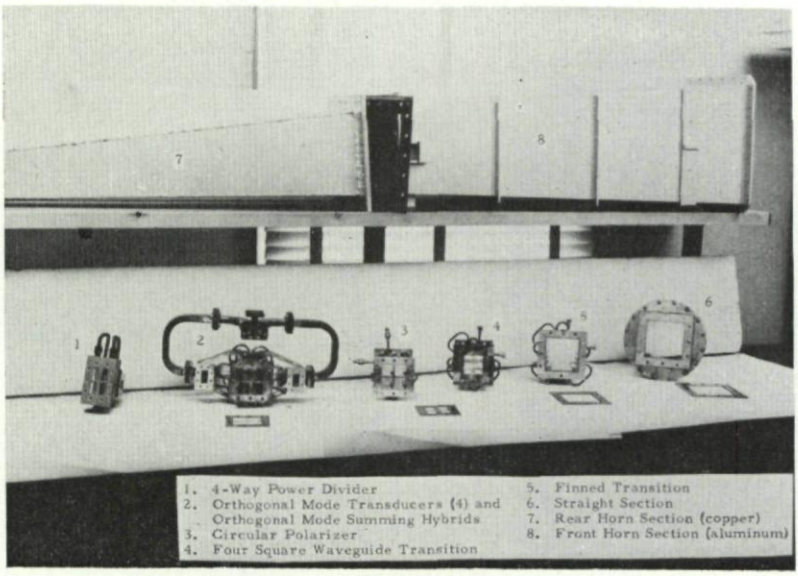


FIG. 32 *Multimode horn subassemblies.*

side lobes. The H-plane side lobes are reduced by the elimination of the TE_{30} mode at the horn throat. Typical patterns for the multimode horn are shown in Fig. 33.

The low-noise receivers in the initial high-power plug-in room are two-stage, diode parametric amplifiers capable of being cooled in either liquid nitrogen or liquid helium. Initially, only two channels will be received, the normal circularly polarized return and the orthogonal circularly polarized signal, at either 7750 mc/s or 8350 mc/s. Provision has been made for modification of the RF system to include simultaneous reception on up to seven receiving channels at 7750 and/or 8060 and/or 8350 mc/s.

The initial experiments will be conducted using batch-filled Dewars. Units manufactured by Cryenco, Inc. (Fig. 34) have demonstrated ability to hold helium for 10 hours and nitrogen for 3 days. A recently developed Dewar using super-insulation held helium for 18 hours. Closed-cycle helium

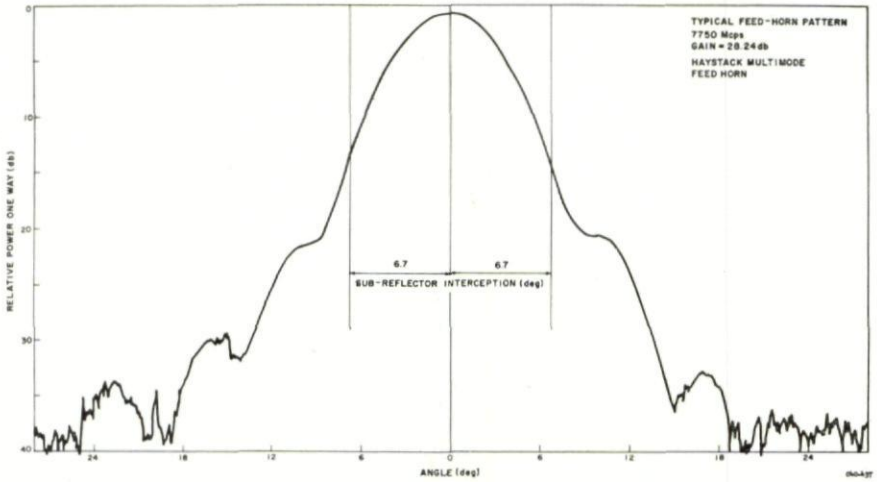


FIG. 33. *Typical pattern of multimode horn.*

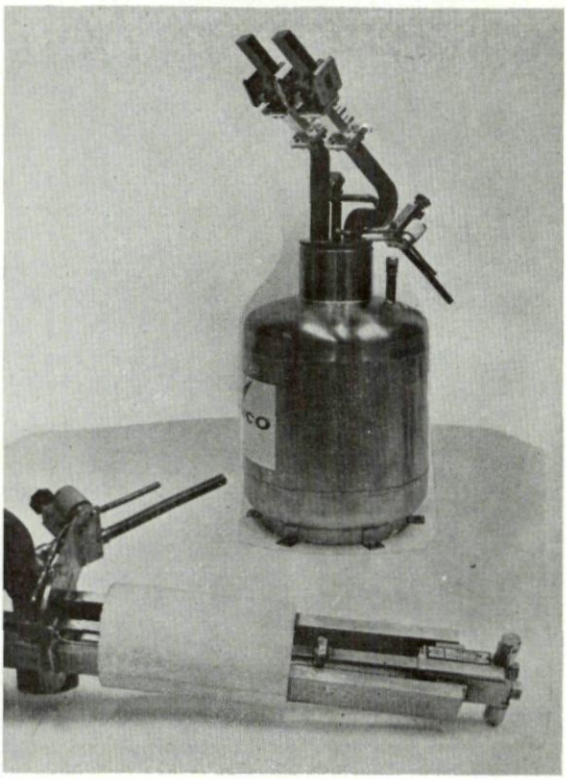


FIG. 34. *Two-stage parametric amplifier and dewar to be used in control tests.*

refrigeration systems (Fig. 35) are undergoing evaluation and will be installed at a later date. Provisions are being made for installation of cryostats in the RF box and compressors on the support structure close to the RF box. Such a system should allow continuous operation at liquid helium temperature. The cryostats have been designed to accommodate four two-stage

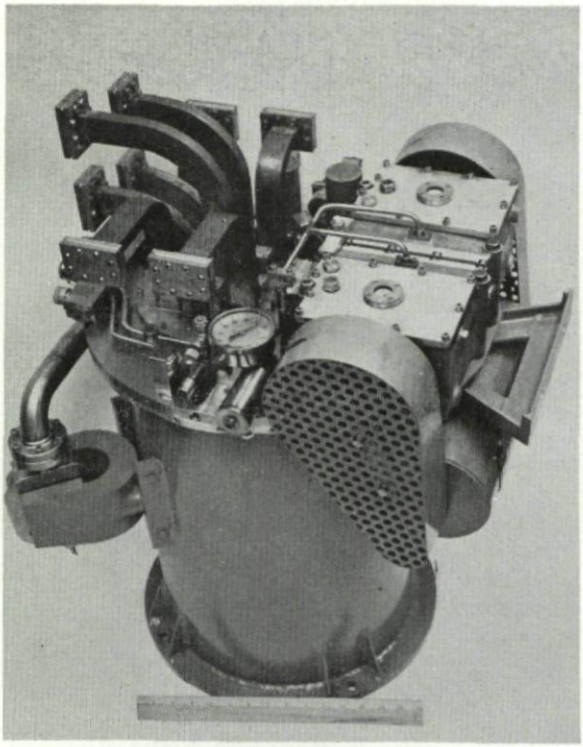


FIG. 35. *A closed Cycle refrigerator and two-stage parametric amplifier.*

paramps, the four first stages being at liquid helium temperature and the second stages located at an intermediate heat station (approximately 60°K).

The parametric amplifier stages to be used in the initial tests are modified Texas Instruments, Inc. units which were originally designed for room-temperature operation. These amplifiers will eventually be replaced by models with lower heat capacity and improved microwave characteristics.

10. CURRENT STATUS

During the month of August 1964, the precision antenna, which has been the pacing item on the over-all HAYSTACK schedule, completed nearly all of its mechanical acceptance tests. The other elements of the HAYSTACK system, including the pointing control computer, the plug-in radiometer instrumentation, the power supply, and the frequency control and timing

equipment, have been completed and are installed at the HAYSTACK site. The antenna pointing and control system has successfully operated a remote antenna over telephone lines, and the computer programs appear satisfactory.

In September 1964, measurements of the gain and noise temperature of the composite antenna-radome configuration will be initiated. These tests will make use of the radiometer equipment and radio star noise sources. The results of these measurements will be compared with data recently taken with the same radiometer and a reference horn antenna to determine the effective gain and noise temperature of the HAYSTACK system.

If these tests are successful, the high-power transmitter will be used for active and passive communications and radar experiments before the end of 1964.

ACKNOWLEDGEMENTS

The planning, design, installation, and testing of this new facility has been successful only because of the hard work, competence and diverse talents of a large number of individuals and organizations. It is not feasible to enumerate the contributions made by each participant in the program during its four-year history.

DISCUSSION

E. RECHTIN: J.P.L. has used the M.I.T. computor programs for antenna design and strongly recommends other antenna designers to investigate this approach. J.P.L. has found it to save time and money and to reduce the risk of design.

CHAPTER 20

THE DESIGN OF A VERY HIGH POWER, VERY LOW NOISE CASSEGRAIN FEED SYSTEM FOR A PLANETARY RADAR*

P. D. POTTER
Jet Propulsion Laboratory, Pasadena, California

A modified Cassegrain feed for an 85 foot diameter antenna, planetary radar system is described. The equipment and techniques involved are being developed for use in the NASA/JPL Deep Space Instrumentation Facility. The system operates at a frequency of 2400 mc/s in a duplex mode with a transmitter power of 100 kW CW and an overall receiving system noise temperature of about 28°K. The feed system utilizes a non-optical sub-reflector and a suppressed sidelobe feedhorn to achieve an overall aperture efficiency of 0.65 and, including atmospheric effects, a zenith antenna noise temperature of 10°K. A versatile polarizing system provides for either sense of circular polarization or any orientation of linear polarization. Detailed design and performance information is presented.

1. INTRODUCTION

We have been concerned with development of equipments and techniques for application in the NASA/JPL Deep Space Instrumentation Facility (DSIF). As an aid to this development activity we have worked with an experimental planetary radar system which involves most of the equipments and techniques required in a deep space communications ground station; in particular large, efficient, low noise ground station antennas are common to both technologies.

The capability of a radar system is directly proportional to the square of the antenna gain and inversely proportional to the receiving system noise temperature; thus the antenna gain should be maximized and the antenna effective noise temperature minimized. At one time it was felt that these two requirements were in conflict;[1] more recently, it has been realized that paraboloidal antenna feed systems can be designed which tend to satisfy both the gain and noise temperature requirements simultaneously.

For paraboloidal antenna systems of the 85 foot or larger class, it is readily shown that simple economic considerations dictate the use of a highly

* This paper presents the results of one phase of research carried out at the Jet Propulsion Laboratory, California Institute of Technology, under Contract No. NAS 7–100, sponsored by the National Aeronautics and Space Administration.

specialized, high performance feed system. The following empirical relationship may be used to estimate the structural cost of a large antenna as a function of its diameter:

$$C = aD^b \tag{1}$$

where C = antenna cost
D = antenna diameter
a = a constant, dependent on design, required accuracy, place of manufacture, etc.
b = a constant ≈ 2.6

The antenna gain, G, is given by

$$G = \pi^2 \eta \left(\frac{D}{\lambda}\right)^2 \tag{2}$$

where η = aperture efficiency
λ = operating wavelength, assumed to be a constant, fixed by system considerations

By combining Eqs. (1) and (2) and taking the total differential of G, it is readily shown that to obtain a specified value of G, the cost value ΔC, of an aperture efficiency increase, $\Delta\eta$, is given by

$$\Delta C \approx \frac{b}{2} \frac{C}{\eta} \Delta\eta \tag{3}$$

For typical values of $\eta = 0.65$ and $b = 2.6$, the fractional cost is:

$$\frac{\Delta C}{C} = 2\Delta\eta \tag{4}$$

Equation 4 means that a 10% increase of aperture efficiency (from 0.60 to 0.70) provides the performance equivalent of reflector diameter change which would result in a 20% increase in structural cost. Since precision antennas of the 85 foot class generally cost $300,000 to $1,000,000 from an economic standpoint alone, the feed system design is worthy of considerable engineering effort.

Also, the reproduction cost of a high efficiency feed is not much greater than a low efficiency feed once the development work is done. However, the cost of reproducing a larger reflector antenna remains a fixed amount greater than reproducing a smaller one. Therefore, economy in duplicating antenna installations will further favor the appropriate expenditure of effort on feed system design.

Assuming the paraboloidal reflector surface is sufficiently accurate for the frequency of operation, all techniques for maximizing the aperture efficiency and minimizing noise temperature basically involve control of the reflector illumination function. This in turn requires control of the feed system radiation pattern and possibly control of the feed system support structure design to minimize aperture blockage effects. Figure 1 shows three idealized

feed system radiation patterns. Figure 1a is the optimum but physically unrealizable case where the paraboloid is uniformly illuminated with no spillover, resulting in maximum aperture efficiency (100%) and minimum noise temperature. Figure 1b depicts the type of illumination achieved with

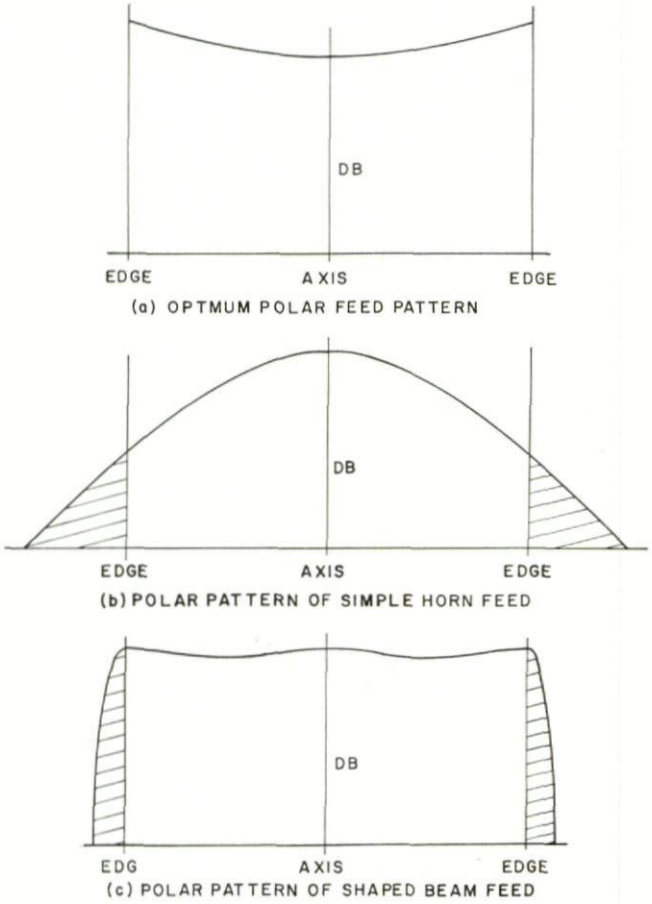


FIG. 1. *Idealized feed patterns.*

a simple horn-type feed. Such a system will typically have an overall aperture efficiency of 0.50–0.60, spillover of 10–20% and a zenith noise temperature contribution of 20°–40°K. Figure 1c indicates a realizable radiation pattern which will yield a high performance level.

Although performance of the general type depicted in Fig. 1c has been realized in focal point feed systems,[2] we have found the two reflector Cassegrain-type configuration to be a more practical and operationally convenient approach. The general physical layout of such a system is shown in Fig. 2. A subreflector is interposed between the focal point of the paraboloid

and the feedhorn center of radiation which, when viewed as operating in the transmit mode, transforms the feedhorn radiation into a quasi-spherical wave centered at the paraboloid focal point; in the limit of vanishing wavelength, the subreflector is a truncated hyperboloid. A Cassegrain-type system in which the subreflector shape deviates from an hyperboloid is referred to as non-optical and generally can be designed to have superior performance relative to a zero wavelength approximation design.

For large ground antennas, the feed support cone may be sizable without significant effect on performance and thus serve as a readily accessible,

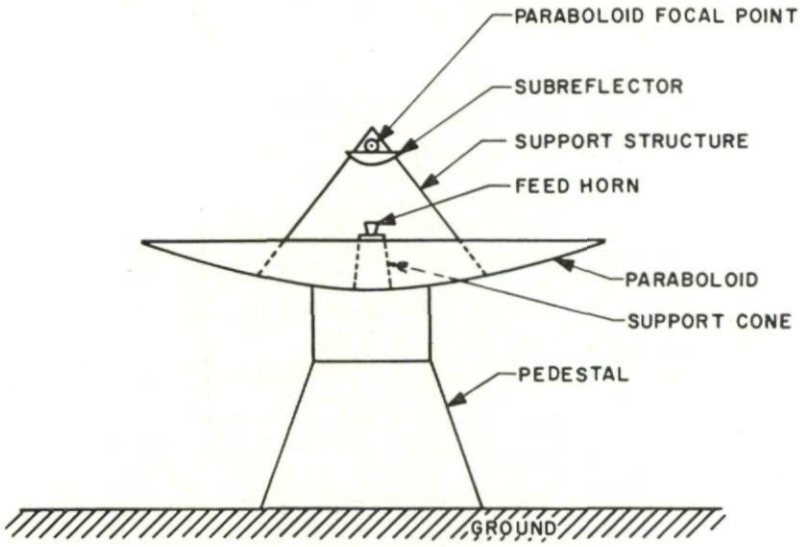


FIG. 2. *Cassegrain antenna.*

environmentally-controlled location for various associated RF equipments. Primarily for these practical reasons and also because of a high performance potential, the Jet Propulsion Laboratory (JPL) feed system was chosen to be a modified Cassegrainian. Operational experience obtained since the system was put into service in October, 1962, has strongly supported the correctness of this choice.

2. PREVIOUS APPLICABLE CASSEGRAIN WORK

The two reflector system invented by W. Cassegrain has been used extensively for many years in optical telescopes, primarily to achieve a long effective focal length with a convenient physical configuration. During the last decade, widespread interest has developed in the use of this type of system at microwave frequencies. An excellent tutorial paper on the microwave application was published by Hannan[3] in 1961. Hannan discusses the Cassegrain antenna primarily from the standpoint of geometric optics and

mentions as possible advantages over a focal point feed system; superior physical configuration, more flexibility in feed system design and the possibility of long equivalent focal length for simultaneous lobing applications.

The specific advantage of the two-reflector system for applications requiring minimum rear hemisphere radiation and high front-to-back ratio was pointed out by Foldes and Komlos[4] in 1960. Detailed experimental measurements are discussed in their paper relating to the deviation of the feed system performance from that which would be expected based on geometric optics; their design was in fact, experimentally optimized taking into account the empirically determined diffraction effects of the (optical) subreflector. Under JPL sponsorship, this work was extended by Foldes to consider the application to large low noise antennas.[5] Almost simultaneously, experimental work at JPL with a low-noise 85 foot diameter Cassegrain system was reported by Potter.[6] This system designed for use at 960 mc/s, made use of an empirically-derived non-optical subreflector[7] for increased performance; the performance reported was a zenith noise temperature of 9.5°K, including atmospheric and extra atmospheric effects (approximately 3.5°K at 960 mc/s), and an aperture efficiency of 0.50.

The microwave performance of a Cassegrain subreflector has been analyzed in detail in a recent paper by Rusch,[8] who points out that the noise temperature performance of the two reflector antenna system is primarily a function of the subreflector diffraction characteristics, i.e., its non-optical behavior. Rusch uses the current distribution method[9] of vector diffraction theory to develop a machine programme for accurate calculation of the scattering pattern of an optically-designed subreflector, and hence the antenna noise temperature contribution due to rearward spillover. This analytical technique has also been extended[10] to consider non-optical subreflectors.

Although the zenith noise temperature performance of a Cassegrain antenna is largely a matter of the subreflector design and wavelength size the aperture efficiency is primarily a function of the feedhorn design, i.e., the manner in which the subreflector is illuminated. For maximum aperture efficiency, the subreflector should be illuminated essentially uniformly by a feedhorn whose radiation pattern is axially symmetric and cuts off sharply at the subreflector edge. Spillover around the sub-reflector edge is a serious problem not only because of aperture efficiency degradation but also because this energy creates a spurious forward sidelobe level, making the antenna susceptible to solar jamming and increasing the antenna temperature at low elevation angles.

For single frequency channel applications such as planetary radar studies, the dual mode conical feedhorn[11,12] represents a significant improvement over conventional feedhorns because of its pattern symmetry and low sidelobe level.

3. FEED SYSTEM CONFIGURATION

The antenna performance as a function of the Cassegrain optics configuration has been considered in detail previously.[5,6] Within a rather broad

range of parameters, aperture efficiency and zenith noise temperature are not particularly sensitive to feedhorn location, subreflector size and other resulting configuration parameters.

The 2400 mc/s planetary radar feed system configuration is shown in Fig. 3. Figure 4 is a photograph of the antenna with feed system installed; a second experimental cone is on the ground nearby. The support cone is designed in such a way that it may be mounted or removed from the antenna in 2–4 hours, thereby permitting extensive system checkout of the enclosed

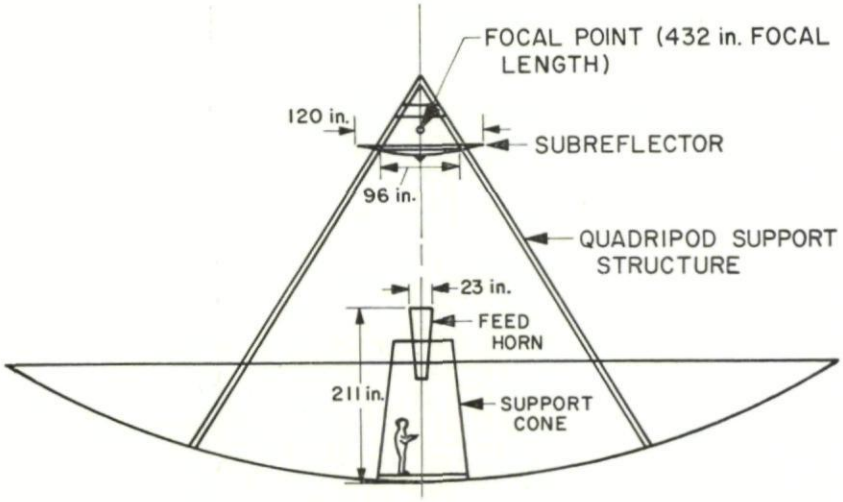


FIG. 3. *Planetary radar feed system configuration.*

receiver and calibration equipment on the ground prior to installation in the antenna (see Fig. 4).

The subreflector of the planetary radar feed system consists of a 96 inch diameter truncated hyperboloid with a 120 inch diameter non-optical beamshaping flange to reduce antenna noise temperature.[7] Reflection of energy by the subreflector back into the feedhorn, with the attendant impedance mismatch, is precluded by use of a centrally located vertex matching plate on the subreflector.[4] The entire subreflector is mounted on a motor driven jackscrew, threepoint support system, thereby allowing for remote focussing of the feed system from the control room. Boresight alignment of the antenna is accomplished by manual adjustment of the three jackscrew supports.

A pictorial block diagram of the support cone equipment is shown in Fig. 5. The polarizer is a turnstile junction type.[13] This type of junction is a six port device; two ports are spatially orthogonal H_{11} circular waveguide modes, two are H_{10} rectangular waveguide outputs and the final two are short circuit terminated H_{10} retangular waveguide ports. By appropriate choice of the short circuit lengths it is possible to excite the feedhorn with any type of polarization. Normally, circular polarization is used for radar

experiments; by manual change of the short circuits, two continuously rotatable spatially orthogonal modes of linear polarization may be obtained. The polarization switch allows remotely controlled choice between right or left-hand circular polarization, or between the two orthogonal linear polarizations.

As shown in Fig. 5, a second waveguide switch is used to switch the polarizer either to the high power transmitter or the receiving system. The third

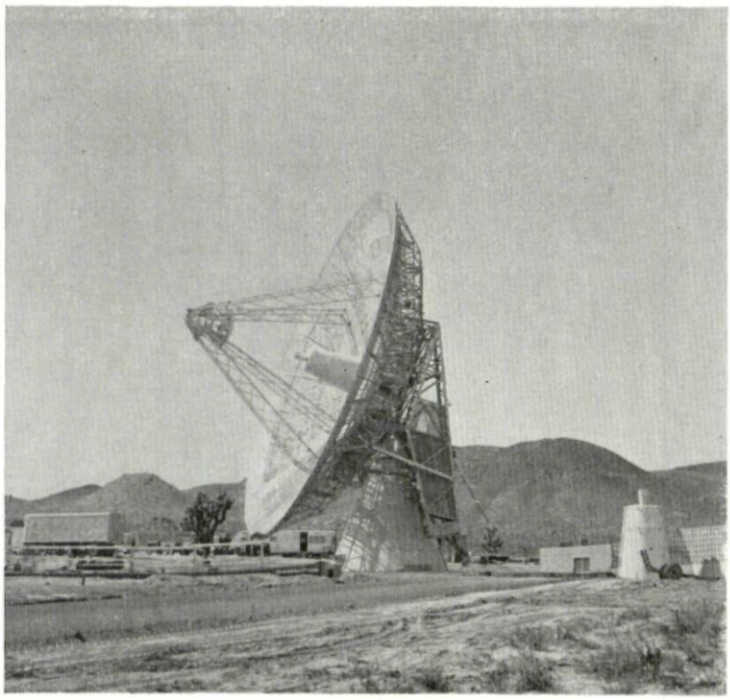


FIG. 4. *Planetary radar antenna system.*

switch allows the receiver (Maser) input to be switched between the antenna or either of two calibrating cryogenic terminations. During the normal transmit mode of radar operation, the polarization switch is in the right circular position, the transmit-receive switch is in the transmit position, and the calibration switch is in the nitrogen load position; the latter to provide additional isolation between transmitter output and receiver input. In the normal receiving configuration, the polarization and transmit-receive switch positions are reversed and the calibrate switch positioned to the antenna port. During radar operation the transmitter drive and switch positions are changed remotely and automatically at time intervals which correspond to the round-trip propagation time between Earth and the planetary target.

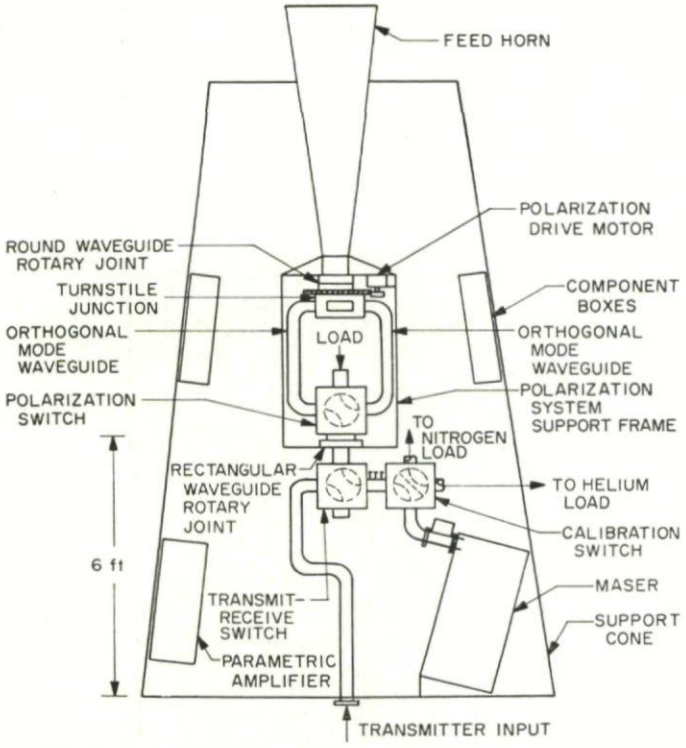


FIG. 5. *Support cone equipment layout.*

The polarization flexibility and excellent axial ratio of the overall antenna system have been employed by Schuster and Levy[14,15] to perform a number of interesting polarization experiments with the planet Venus as a radar target.

4. MICROWAVE COMPONENT DESIGN

The feedhorn is a dual mode conical horn;[11,12] this type of horn is basically a small flare angle conical horn excited with both the dominant H_{11} and higher-order E_{11} cylindrical waveguide modes, in such a way that the resulting radiation pattern is essentially axially symmetric, has low sidelobes and negligible cross polarization. The measured feedhorn radiation patterns in the electric, magnetic and diagonal planes are shown in Figs. 6, 7, and 8 respectively. Because of the desired presence of the higher-order E_{11} mode in the horn, it is necessary to maintain extreme mechanical precision (the order of 10^{-3} wavelength) on the inner surfaces to prevent mode conversion, with attendant pattern degradation. For this reason very conservative mechanical design and fabrication techniques were used; the entire feedhorn was machined from three solid aluminum billets; a typical section is shown in Fig. 9. The completed feedhorn undergoing radiation pattern

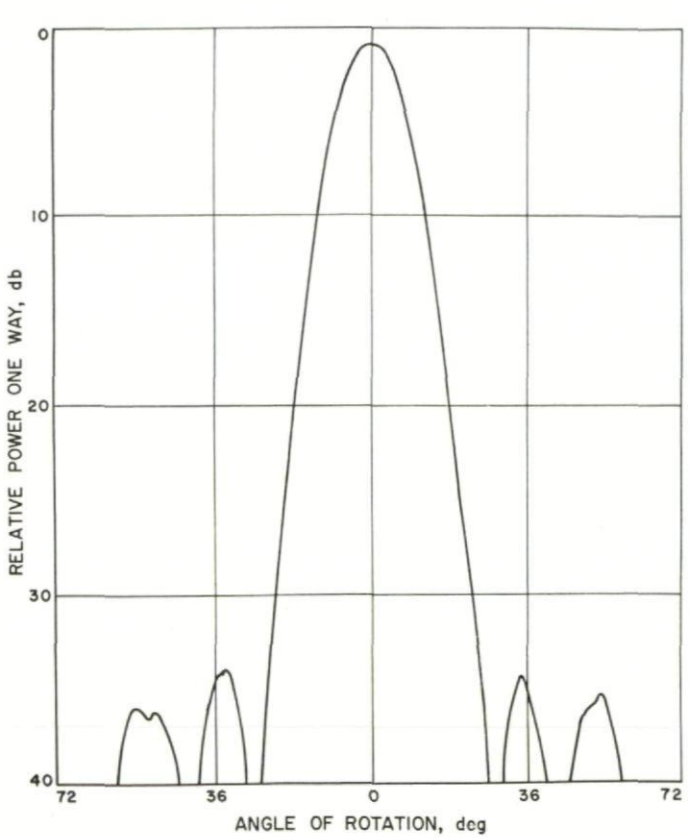


FIG. 6. *Feedhorn radiation pattern, E-plane.*

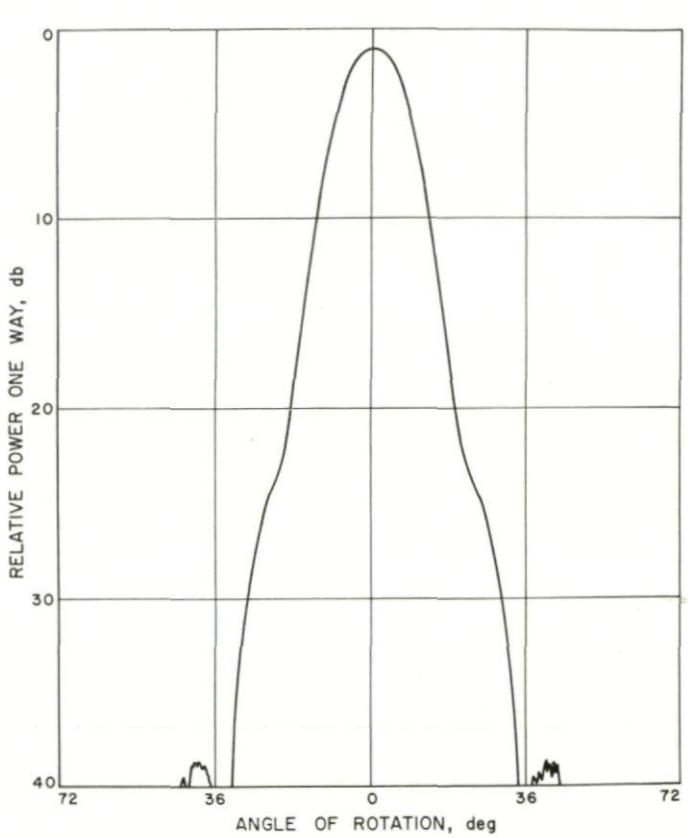


FIG. 7. *Feedhorn radiation pattern, H-plane.*

tests is shown in Fig. 10. It is entirely a figure of revolution allowing complete polarization flexibility.

The turnstile junction polarizer presented a special problem because of the required high power level (100 kW, CW) during transmit operation. A photograph of the unit is shown in Fig. 11. The cylindrical device protruding from the circular waveguide output is the tuning probe, required

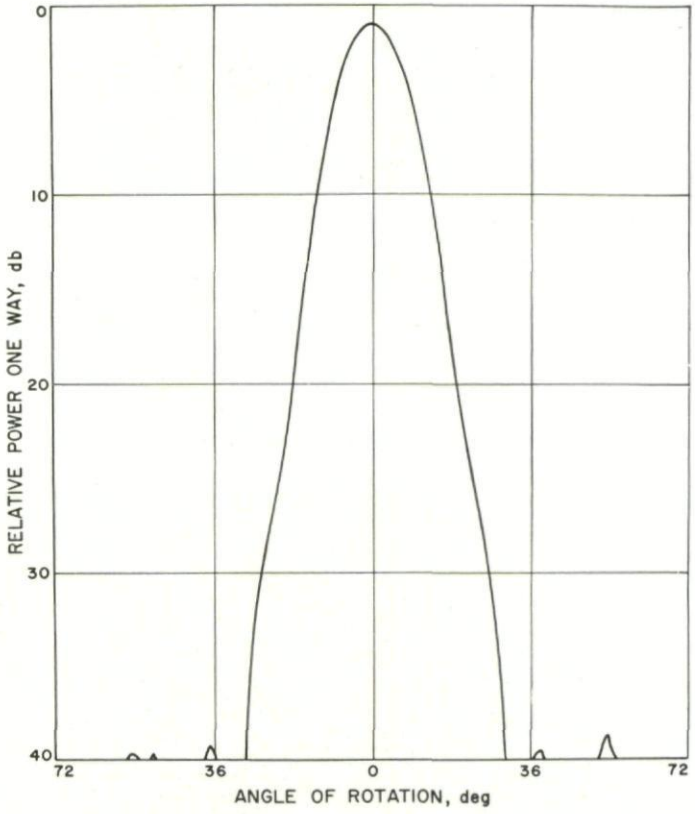


FIG. 8. *Feedhorn radiation pattern, 45° plane.*

for mismatch elimination and port isolation. Between this probe and the circular waveguide wall, TEM coaxial mode energy exists which may cause excessive field strength at the probe surface, with possible arcing. High power tests of the turnstile junction assembly were utilized, in conjunction with conventional microwave measurements, to determine a matching probe configuration which had the desired electrical performance and at the same time would operate at the 100 kW power level. The ellipticity produced by the polarizer, as measured on the entire antenna system is 0.4 db.

The waveguide rotary joints, switches, bends and other components are of conventional design and did not present any particular problem. All

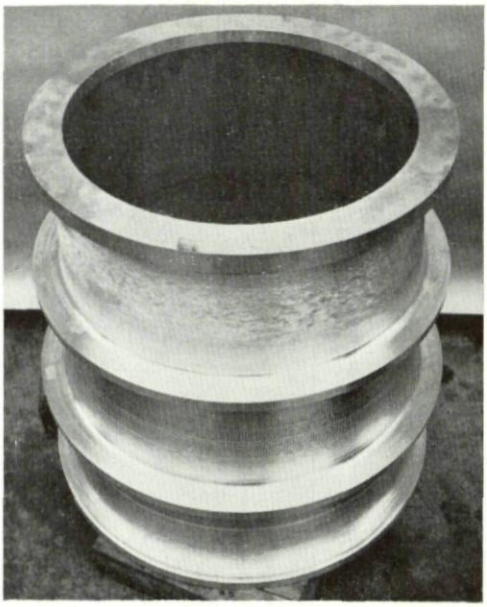


FIG. 9. *Feedhorn section after rough machine.*

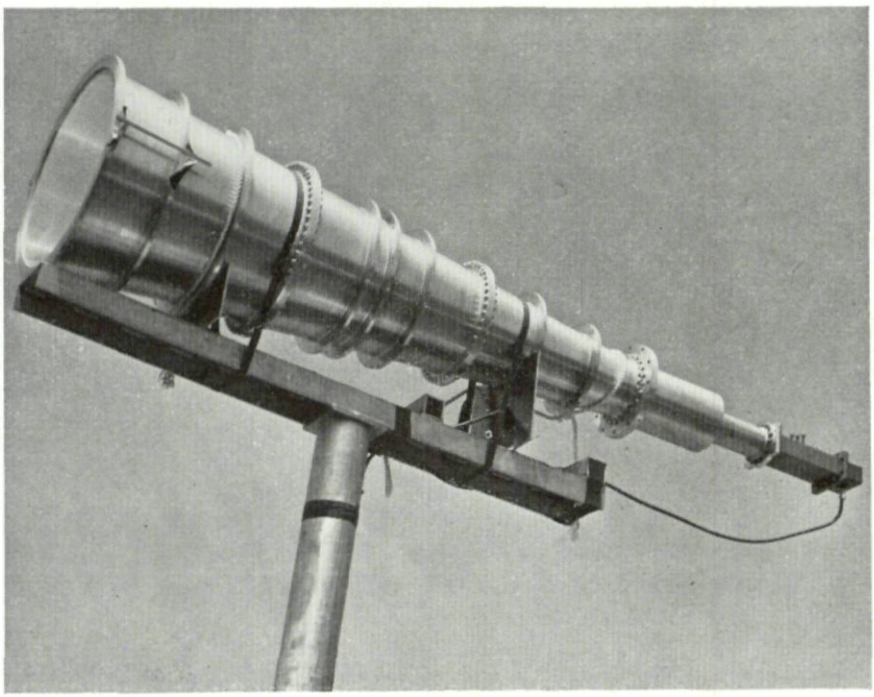


FIG. 10. *Feedhorn undergoing test.*

mating flanges were, however, carefully machined and hand-lapped flat. All components are individually matched to a voltage standing wave ratio (VSWR) of 1.01 to 1.05, and inside surfaces were carefully cleaned to reduce loss and prevent breakdown under power. All high power waveguide parts except the switches and those which are in the rotating assembly are water cooled for temperature stability and the entire waveguide system including

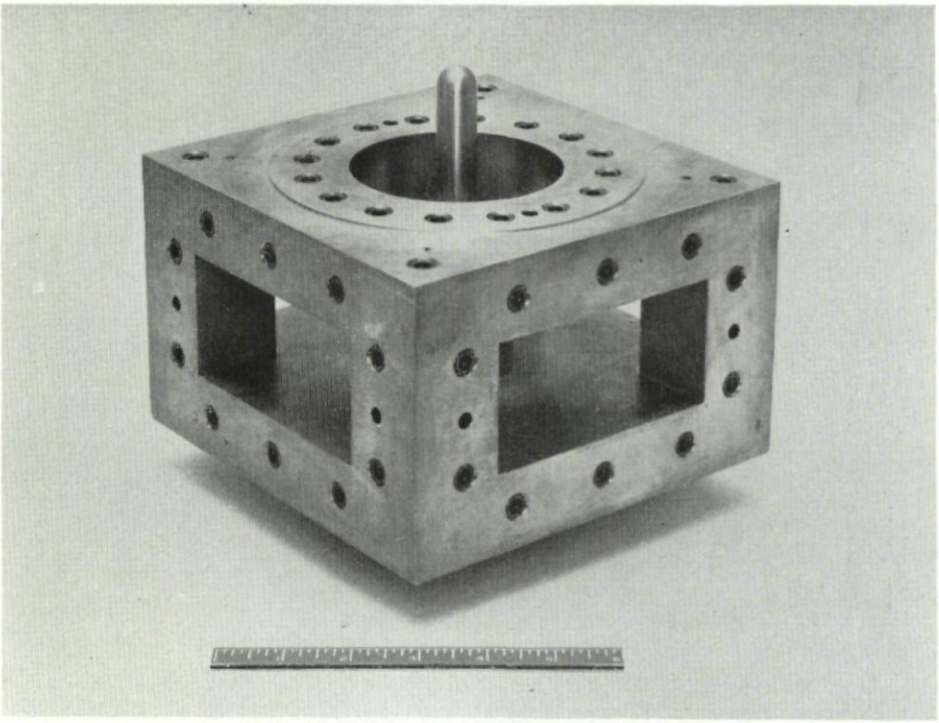


FIG. 11. *High power turnstile junction.*

the feedhorn is pressurized with dry nitrogen at 0.2 psi above atmospheric pressure to prevent corrosion and to insure an inert gas for the transmitter tube should its vacuum seal leak.

5. GAIN AND NOISE TEMPERATURE PERFORMANCE

The accurate gain calibration of large antenna systems poses a special problem because of the large distance involved in far field tests. For an 85 foot antenna operating at 2400 mc/s, the conventional "far-field" distance, i.e., the distance given by twice the ratio of antenna diameter squared to wavelength, is 6.6 statute miles, and an even larger distance is desired for accurate gain calibrations. As the solution to this problem, a calibration installation was made on a mountain, Mt. Tiefort, 12.7 miles distant from

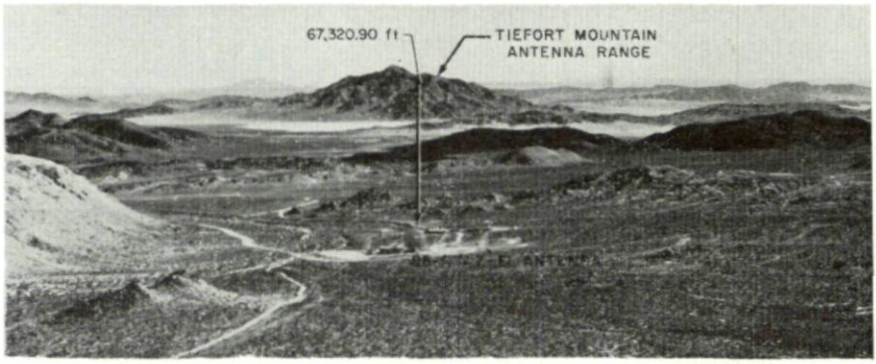


FIG. 12. *Tiefort mountain antenna range.*

the radar antenna. Figure 12 is a photograph showing the radar installation, Mt. Tiefort and the resulting measurement range.

The technique used for gain calibration involved direct measurement of the signal attenuation between the antenna and the Mt. Tiefort signal source, using suitable corrections for atmospheric loss due to oxygen molecule resonance.[16,17] The equipment block diagram for the test is shown in Fig. 13. A standard deviation of 0.15 db was obtained for the test as shown in Table 1 below.

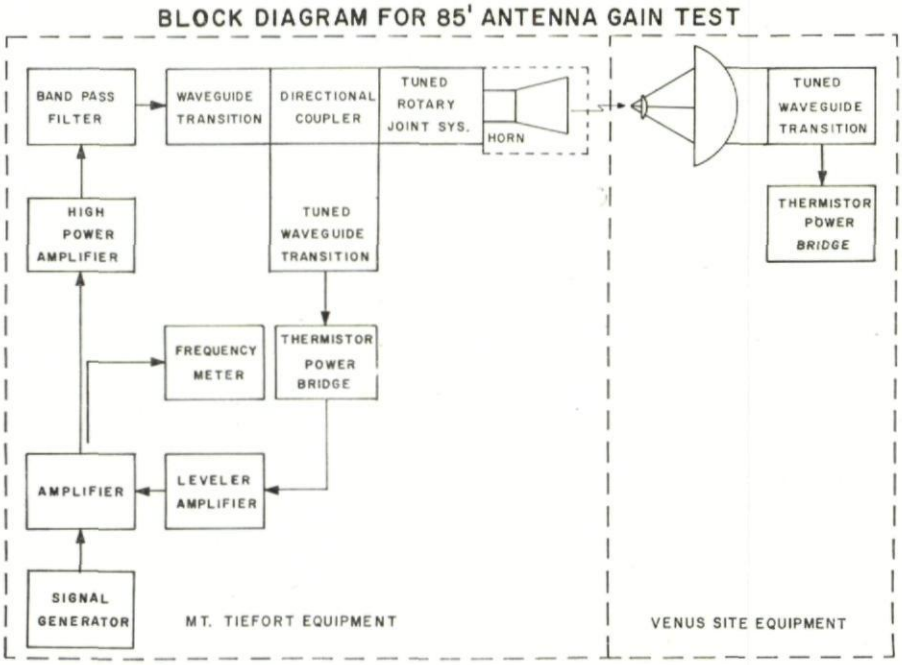


FIG. 13. *Block diagram for 85′ antenna gain test.*

Usually, uncertainty in the standard horn calibration is the most serious error in this type of test. For this reason, a unit identical to the feedhorn was calibrated and used; for the standard gain application, this type of horn has the multiple advantages of: low sidelobes, negligible cross polarized radiation, a well defined phase center close to the aperture, equal beamwidths in all planes, and complete polarization versatility. The horn gain was determined by three independent methods: (1) direct measurement, using a pair of identical units; (2) numerical integration of measured radiation patterns

TABLE 1
ANTENNA GAIN CALIBRATION ERRORS

Error	Estimated standard deviation in decibels (db)
Standard Horn Gain	0.09
Signal Attenuation	0.05
Multipath	0.05
Coupler Calibration	0.05
Atmospheric Loss	0.05
Atmospheric Mismatch	0.05
Dissipative Losses	0.03
Antenna Alignment	0.02
Range Length	0.01
Total Standard Deviation	0.15 db

and (3) analytical integration of the predicted aperture distribution. These three techniques yielded values, respectively, of 22.17 db, 21.95 db, and 21.85 db. A statistical analysis of these data and the associated errors yields an expected gain value of 22.02 db with a standard deviation of .09 db.

Using the above results, the 85 foot antenna gain at the feedhorn output was found to be 54.40 db above isotropic, with a standard deviation of 0.15 db; the corresponding aperture efficiency is 0.65. This measurement is felt to be one of the most accurate gain calibrations of an 85 foot antenna ever performed and is presently being used[18] for calibration of radio source absolute flux density. This work, when completed, will provide a convenient and accurate means of gain evaluation and comparison of the antennas in the NASA/JPL Deep Space Instrumentation Facility.

General techniques for a determination of effective antenna temperature have been derived by Schuster et al.[2] The method used for evaluation of the planetary radar system basically consists of comparing the noise power received by the antenna with that emitted by the liquid nitrogen and liquid helium terminations. Corrections for the small (0.1 db) insertion losses in the various transmission paths were made, resulting in an overall standard deviation for the zenith antenna temperature of approximately 0.75 °K. The mean value for the antenna temperature at the feedhorn output is 10.5 °K. Figure 14 is a plot of the antenna temperature as a function of elevation angle, together with the temperature which would be observed if the only

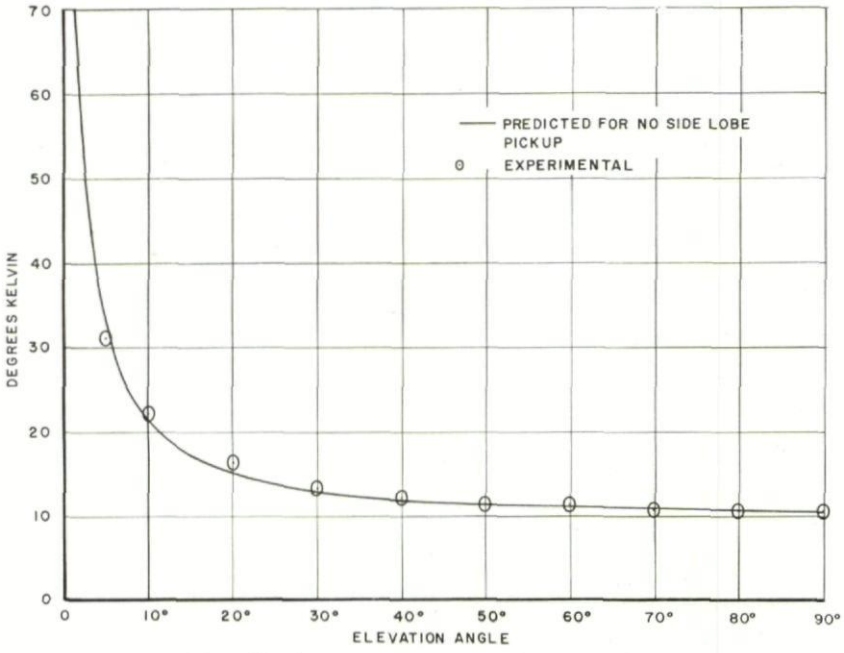


FIG. 14. *Antenna temperature vs. elevation angle.*

elevation angle dependence were that predicted by Hogg[17] for the atmosphere. It should be noted that the forward sidelobe contribution to antenna noise temperature is scarcely discernible.

6. PERFORMANCE ANALYSIS AND THE POSSIBILITY OF FUTURE IMPROVEMENT

The radiation pattern of a paraboloidal antenna may be calculated by integration of its aperture distribution;[19] the radiation pattern, $g(\theta, \phi)$, is given by

$$g(\theta, \phi) = \int_0^{2\pi} \int_0^a F(\rho, \phi')\, e^{jk\rho \sin\theta \cos(\phi - \phi')} \rho \, d\rho \, d\phi' \qquad (5)$$

where θ = antenna polar pattern angle
ϕ = antenna azimuthal pattern angle
k = free space propagation constant
ϕ' = azimuthal coordinate of a point in the aperture
ρ = radial coordinate of a point in the aperture
a = aperture radius
$F(\rho, \phi')$ = complex aperture field distribution

Equation (5) is based on scalar diffraction theory and is valid when $F(\rho, \phi')$ is a slowly varying function of position and has small phase error, and when

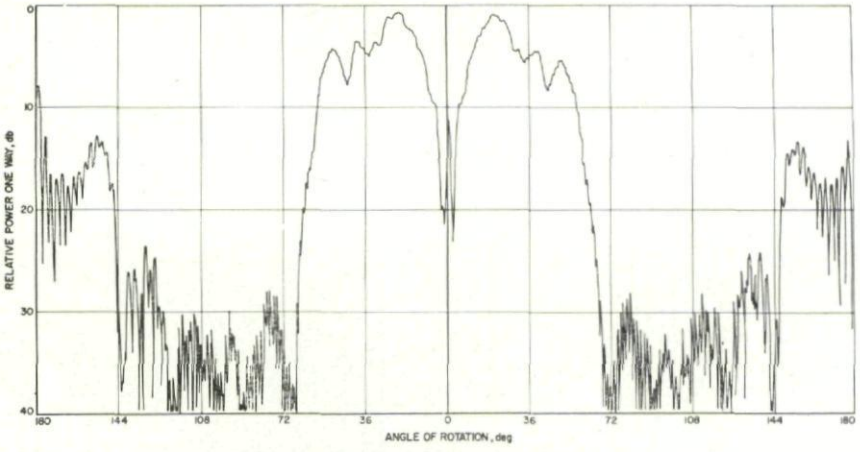


FIG. 15. *Feed system radiation pattern.*

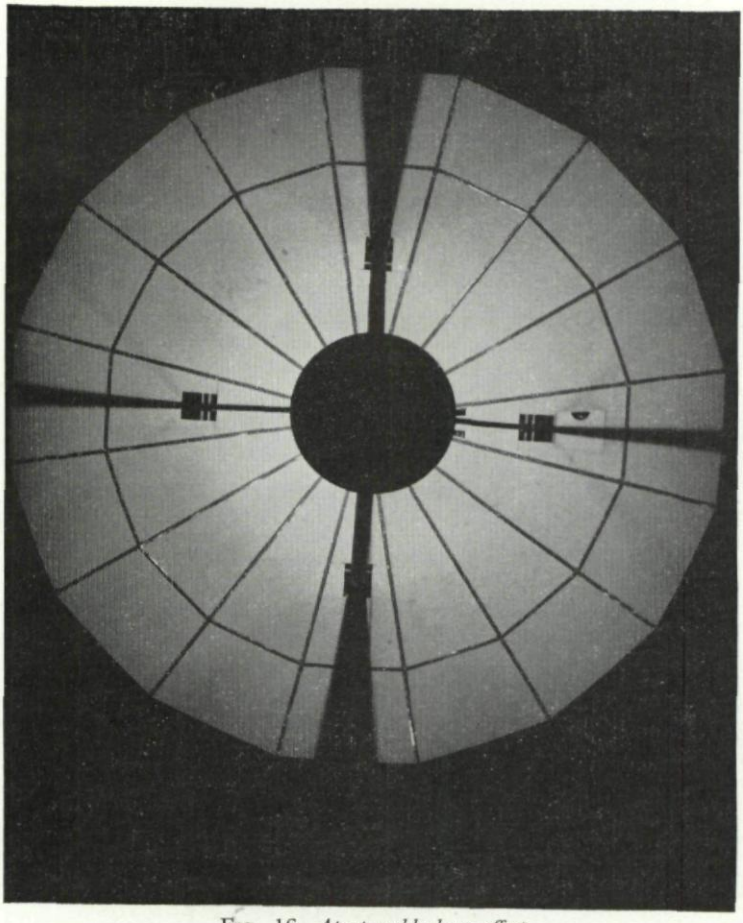


FIG. 16. *Aperture blockage effect.*

θ is much smaller than 90°. $F(\rho, \phi')$ is established by two factors: the feed system radiation pattern and the aperture blockage.

The planetary radar feed system polar radiation pattern, i.e., the pattern of the feedhorn-subreflector combination, is shown in Fig. 15. The relatively uniform radiation in the angular region subtended by the paraboloid ($\pm 60°$)

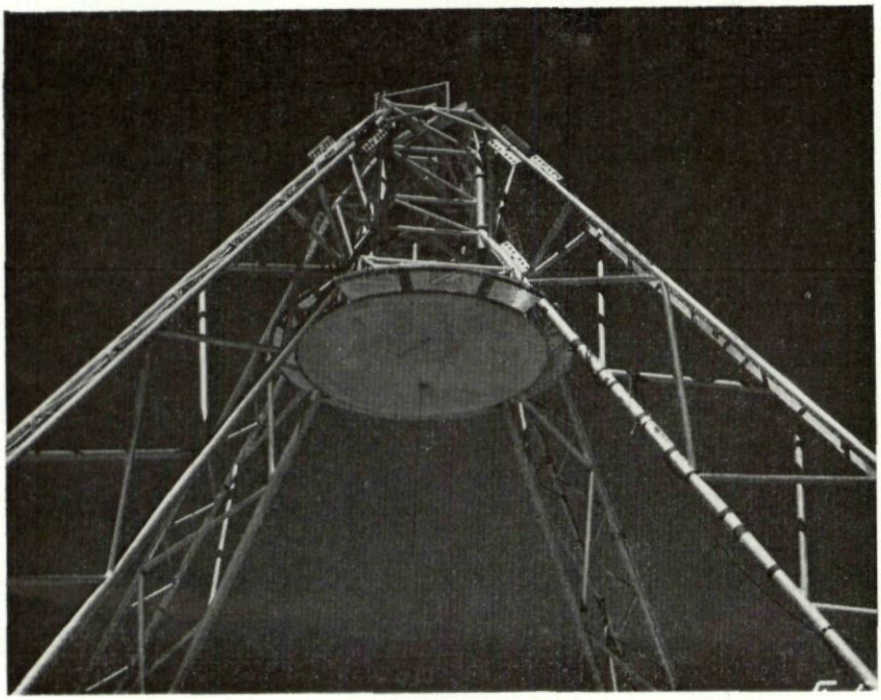


FIG. 17. *Truss type quadripod.*

and low rearward spillover are evident. The radiation pattern, experimentally determined by scale model tests and also by machine computation[10] is essentially axially symmetric, as would be expected from the symmetric feedhorn pattern and subreflector of revolution.

Typical aperture blockage effects are depicted in Fig. 16, which is a nighttime photograph taken of a smaller Cassegrain antenna in which the feedhorn had been replaced by a suitable light source. The blockage is of two types: (1) a central circular area caused by the subreflector, and (2) radial wedge-shaped areas caused by the quadripod support structure. The physical configuration of the quadripod support and the subreflector on the 85 foot radar antenna may be seen in Fig. 17.

A machine program has been developed at JPL for evaluation of Eq. (5), using feed radiation pattern data and calculated aperture blockage. The quadripod is included in the calculation as four wedge-shaped regions,

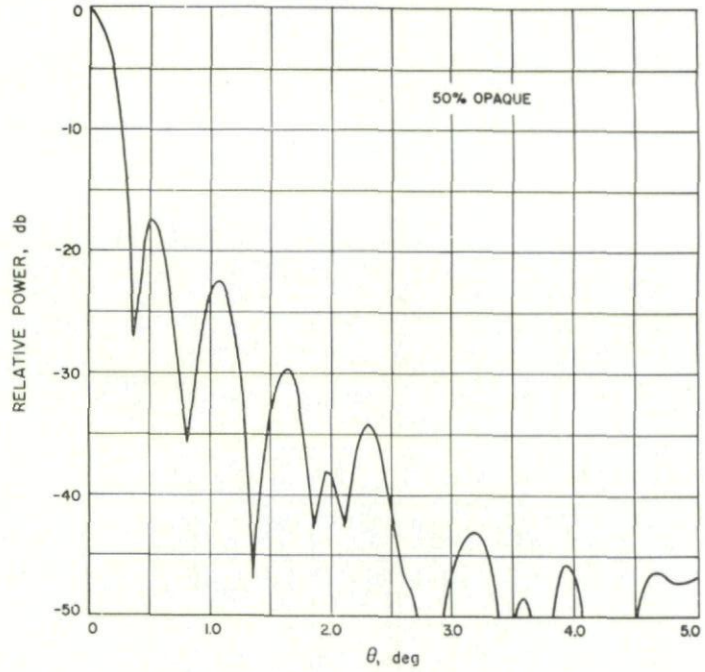


FIG. 18. *Calculated secondary pattern, 50% opaque.*

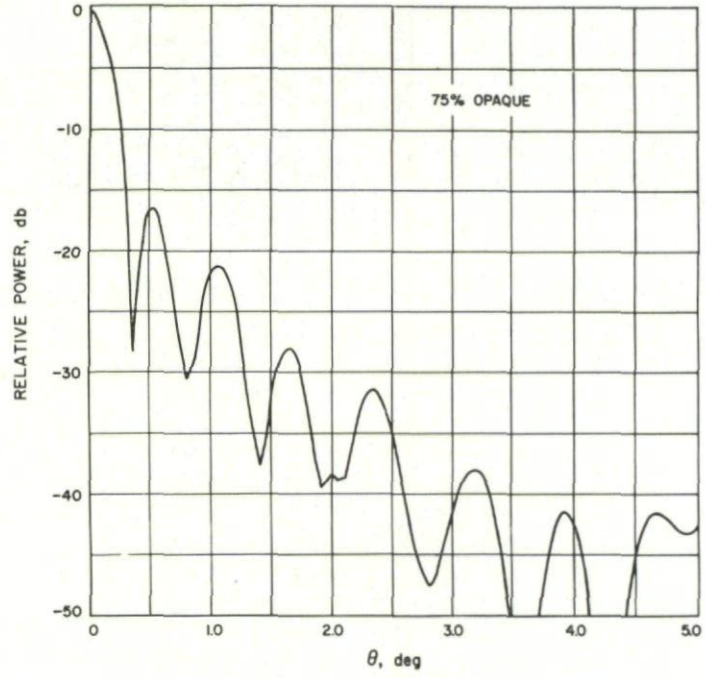


FIG. 19. *Calculated secondary patterns, 75% opaque.*

corresponding to the physical outline of the trusswork. An opaqueness factor is introduced in the quadripod-blocked region to account for the fact that the trusswork will, on average, transmit collimated energy varying from zero (100% opaqueness) to full illumination (0 opaqueness). Figures 18 and 19 are machine computed secondary patterns of the 85 foot antenna for 50% and 75% opaqueness, respectively. The true opaqueness of this quadripod

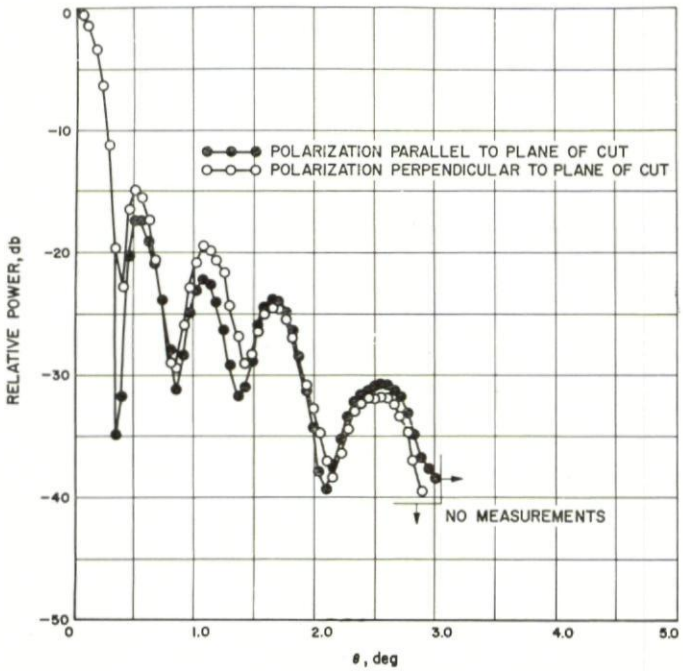


FIG. 20. *Measured secondary pattern.*

is estimated by a comparison with the experiment radiation patterns shown in Fig. 20. Although there is clearly a quadripod polarization dependence, the opaqueness appears to lie between 50% and 75%. From a knowledge of the geometrical outline blockage of the quadripod, which is 14.4%, the quadripod microwave blockage is found to lie between 7% and 11%. The corresponding loss of aperture efficiency due to this effect is −0.6 to −1.0 db. As shown below, a value of −0.73 db (9% blocking) results in perfect agreement between calculated and measured aperture efficiency, and is considered a reasonable estimate of the true effect.

The predicted (calculated) overall gain and aperture efficiency are compared with the measured values in Table 2 below.

The illumination factor of −1.06 db was calculated by numerical integration[20] of the feed system radiation pattern shown in Fig. 15. As mentioned above, there is no significance to the exact value of 63% opaqueness, other than the ensuing agreement between calculated and measured antenna gain.

TABLE 2
PREDICTED AND MEASURED GAIN AND APERTURE EFFICIENCY

Item	Associated gain, db	Aperture efficiency
Theoretical maximum	56.24	1.000
Illumination factor (includes phase loss)	−1.06	—
Gain for perfect surface and no quadripod	+55.18	0.783
Surface tolerance loss (0.032 in. rms)	−0.05	—
Gain for no quadripod	+55.13	0.775
Gain loss for 100% opaque quadripod (machine computed)	−1.19	—
Gain loss for 63% opaque quadripod	−0.73	—
Gain for 63% opaque quadripod	+54.40	0.655
Measured gain	54.40 ± 0.15 s.d.	0.655 ± 0.020 s.d.

Table 2 not only shows that antenna gain may be precalculated to good accuracy, but also indicates possibilities for future improvement. Although the illumination loss is only −1 db, it appears based on present studies in progress at JPL, that further feed system development could effect an improvement of roughly 0.5 db in this area. Similarly, it appears that the effective quadripod blockage could be reduced by more efficient structural

TABLE 3
PREDICTED AND MEASURED ZENITH NOISE TEMPERATURE

Feed Spillover (0.5%)	1.0°K (Predicted from scale model tests)
Quadripod Scattering	5.5° (Predicted from 9% blocking, energy assumed to scatter isotropically, averaged 240°K ground)
Surface Leakage Between Panels	0.5° (Extrapolated from a measured value at a different frequency).
Atmosphere and Extra Atmospheric Noise	3.0 (Measured)
Predicted Total	10.0°K
Measured Total	10.5°K ± 0.75 s.d.

design; reduction of this effect to −0.25 db (3% blockage) appears reasonable. An overall improvement of approximately 1 db, therefore, appears possible.

Predicted and measured zenith antenna noise temperature performance are shown in Table 3 below.

Table 3 demonstrates that zenith noise temperature may be predicted to good accuracy from a knowledge of the feed system patterns and the antenna physical characteristics. Reduction of the quadripod blockage to 3% as described above would reduce that contribution to 2°K; also future feed system improvement to 0.5° spillover contribution appears reasonable. Therefore it is felt that a total zenith noise temperature, at *S*-band frequency, of 6°–7°K is a practical goal for future paraboloidal antenna systems. With no other receiving system improvements, this could reduce the overall system noise temperature from 28°K to 24°K, an improvement in sensitivity of 0.7 db. Improvement would, of course be more dramatic with lower-noise receiving systems.

CONCLUSIONS

A 2400 mc/s Cassegrain-type 85 foot diameter paraboloidal antenna for a planetary radar system has been described. It is applicable in deep space communication systems. It is shown that the performance level is within approximately 2 db of fundamental limits. Based on this fact, economic considerations, and favorable operational experience over a two year period, it is felt that a two-reflector paraboloidal antenna is a sound approach to large aperture, low noise applications such as planetary radar.

ACKNOWLEDGEMENTS

The work reported in this paper results from the efforts of numerous individuals at the Jet Propulsion Laboratory, Pasadena, California. The application of the Cassegrain principle to large, low noise antennas was first suggested to the JPL antenna staff by Dr. E. Rechtin. The planetary radar project was conceived, designed and supported by Messrs. W. K. Victor and R. Stevens. The author is also indebted to the latter for his continuing interest in and support of low-noise systems, particularly antenna feed systems. The polarization system was conceived, designed and tested by Messrs. G. S. Levy, R. Petrie and the late Danver Schuster, who also participated with the author in many stimulating feed system design discussions. The author is also indebted to Messrs. D. Nixon and F. McCrea for assistance in the installation and testing of the system, and to D. Bathker for calculation of the secondary patterns.

DISCUSSION

————: Could Mr. Potter please give us some idea of the bandwidths of his antenna?

P. D. POTTER: I was afraid someone would ask that, the system is actually quite narrow band—the system requirement was for a narrow band antenna.

The major factor which makes it narrow band is the feed horn itself, which as I mentioned has two wave guide modes and which are generated in a tapered region of the horn and which have to be in a certain phase relationship at the aperture. For this reason the band width of this system is approximately 100 Mc/s in 2400.

R. BLOMMENDAAL: How did you design the combination so that efficient illumination resulted?

P. D. POTTER: The feed horn itself is excited in the H_{11} E_{11} wave guide modes. A combination of these two modes has two effects. One is to broaden the pattern in the E plane so that it has a beam width the same as that in the H plane and at the same time it is possible to essentially cancel the side lobe structure of the horn. The shape of the beam in this case is more or less conventional but the loss which is normally incurred due to cross polarisation is essentially eliminated and the spillover efficiency is increased in this case to approximately 96%. In the case of the subreflector the primary problem which is encountered in the Cassegrain system is that the taper of the field in the region which is the boundary between the shadow and light region optically is rather poor at the wavelength sizes which are generally used. The scattering pattern of the subreflector in that region is primarily determined by the surface of the reflector near the outside edge. By an empirical technique we modified the reflector surface in this region to increase the slope of the pattern in the region near the shadow boundary. Since that time we have developed an analytical technique for generating this surface by a machine programme.

REFERENCES

1. CUTLER, C. C., Parabolic-antenna design for microwaves, *Proc. Inst. Radio Engrs., N.Y.*, **35,** 1284–1294, 1947.
2. D. SCHUSTER et al, The determination of noise temperatures of large paraboloidal antennas, *Trans. IRE*, **AP-10,** 3, 286–291, 1962.
3. HANNAN, P. W., Microwave antennas derived from the Cassegrain telescope, *Trans. IRE*, **AP-9,** 140–153, 1961.
4. FOLDES, P., and KOMLOS, S. G., Theoretical and experimental study of wideband paraboloid antenna with central reflector feed, *RCA Review*, **XXI,** 94–116, 1960.
5. FOLDES, P., The capabilities of Cassegrain microwave optics systems for low noise antennas, *Proceedings of the Fifth Agard Avionics Panel Conference, Oslo, Norway*, **4,** 319–325, Pergamon Press, 1962.
6. POTTER, P. D., The application of the Cassegrain principle to ground antennas for space communications, *Trans. IRE*, **SET-8,** 154–158, 1962.
7. POTTER, P. D., Unique feed system improves space antennas, *Electronics*, **35,** 36–40, 1962.
8. RUSCH, W. V. T., Scattering from a hyperboloidal reflector in a Cassegrainian feed system, *Trans. IEEE*, **AP-11,** 4, 414–421, 1963.
9. SILVER, S., *Microwave Antenna Theory and Design*, MIT. Series, Vol. 12, McGraw-Hill, New York, 1949, pp 133–150.
10. RUSCH, W. V. T., Scattering of a spherical wave by an arbitrary truncated surface of revolution, *Jet Propulsion Laboratory Technical Report No.* 32–434, 1963.
11. POTTER, P. D., A new horn antenna with suppressed sidelobes and equal beamwidths, *Microwave J.*, **VI,** 6, 71–78, 1963.

12. POTTER, P. D., and LUDWIG, A. C., Beamshaping by use of higher order modes in conical horns, *Nerem Record*, 92–93, 1963.
13. POTTER, R. S., The analysis and matching of the trimode turnstile waveguide junction, *Naval Res. Lab*, Washington, D.C., Rep. No. 4670, 1955.
14. SCHUSTER, D., and LEVY, G. S., Faraday rotation of Venus radar echoes, the *Astron. J.* **69,** 1, 42–48, 1964.
15. LEVY, G. S. and SCHUSTER, D., Further Venus radar depolarization experiments, the *Astron. J.* **69,** 1, 29–33, 1964.
16. VLECK, J. H. VAN., Absorbtion of microwaves by Oxygen, *Phys. Rev.*, **71,** 6, 413–433, 1947.
17. HOGG, D. C., Effective antenna temperatures due to Oxygen and water vapor in the atmosphere, *J. Appl. Phys*, **30,** 9, 1417–1419, 1959.
18. LEVY, G. S., Jet Propulsion Laboratory, Pasadena, California, U.S.A., private communication.
19. SILVER, S., *Microwave Antenna Theory and Design*, MIT Series (McGraw-Hill, New York), Vol. 12, p 192.
20. POTTER, P. D., Aperture illumination and gain of a Cassegrainian system, *Trans. IEEE* **AP-11**, 3, 373–375, 1963.

CHAPTER 21

ELECTRONICALLY SCANNED ANTENNA SYSTEMS, I

Moses A. Diab
Thomas Maggio
Rome Air Development Center,
Griffiss Air Force Base, New York, U.S.A.

During recent years much effort has been expended in the development of techniques for accomplishing electronic scanning for application to phased array radars. This has led to the development of a variety of beam scanning and beam forming techniques. The purpose of this report is to acquaint the reader with the various forms and implementation of these techniques which have been developed to date. A bibliography of those techniques which have appeared in print is included for those who wish to pursue the subject in greater detail.

1. RESEARCH OBJECTIVE

Recent years have seen the development and implementation of large high power, high gain mechanical antenna systems, the most notable of which are the BMEWS Radars.

Limits are rapidly being reached, however, in both size and power-handling capabilities. Extremely close tolerances must be held in the manufacture of the dishes; waveguides must be highly pressurized and transmitter tubes are reaching their peak power limitations.

For these and other reasons, the phased array radar scanned electronically has received increasing attention over the past few years. Phased arrays are not power limited since the total radiated power is obtained by the coherent addition of a large number of relatively low power sources; thus, the problems of waveguide and transmitter tube power limitations are circumvented. Furthermore, electronic scanning offers the advantage of being able to move the beam to any point in the coverage in an inertialess manner and with a very short response time.

Of course, a whole new set of problems arises in the use of phased array radar, for example mutual coupling between elements and limited angle of coverage. New techniques for extending the coverage and minimizing the mutual coupling problem are presently being investigated.

Perhaps the area that has received the most attention has been methods for achieving the electronic scanning of the beam. A number of novel techniques for achieving electronic scan have been developed and it would be of interest if some of these were briefly reviewed.

It is well known that if we have an array of elements whose excitations are all in-phase there is produced a radiation pattern perpendicular to the plane of the array. If one were now to introduce an incremental phase shift between the elements as shown in Fig. 1, the resultant radiation pattern would be shifted to some angle as indicated. If we were now to devise some

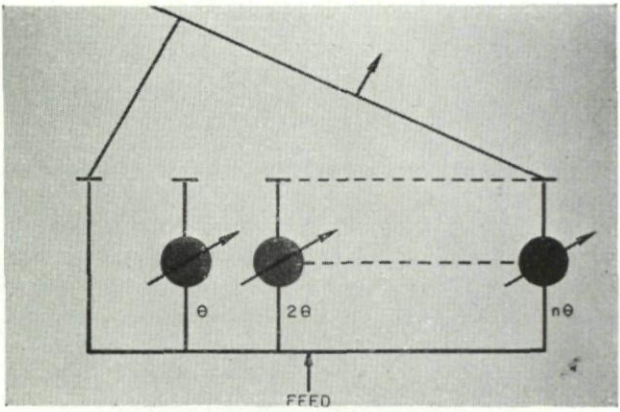


FIG. 1. *Linear array.*

means for continuously varying the phase, it is obvious that the beam can be scanned through some range of angles.

2. TECHNIQUES

The two principle techniques developed and used to date for accomplishing electronic steering of a beam are frequency scan and phase scan. In the frequency scan technique a varying frequency is applied to a length of waveguide which produces a phase gradient along the length of the guide. As the frequency source is varied this phase gradient also varies, producing a scanning of the beam. The energy may then be radiated through slot radiators in the wall of the guide. Two forms of frequency scans are illustrated in Fig. 2.

In a straight waveguide a relatively large frequency excursion is required to move the beam over a small angle. In order to achieve wide angle beam scanning with a reasonable amount of frequency shift, the length of guide between taps must be increased. This has given rise to what has become known as the "serpentine" feed and has found practical application in some radars. Figure 3 shows an Army version of a frequency scan radar originally developed for the US Navy. This radar uses conventional mechanical scan in azimuth but frequency scan in elevation, thereby also providing height information.

Figure 4 shows a system application of the frequency scan technique. The one on the *right* is the AN/SPS-32 search radar which uses frequency scan in one dimension and the one on the *left* is the AN/SPS-33 tracking radar

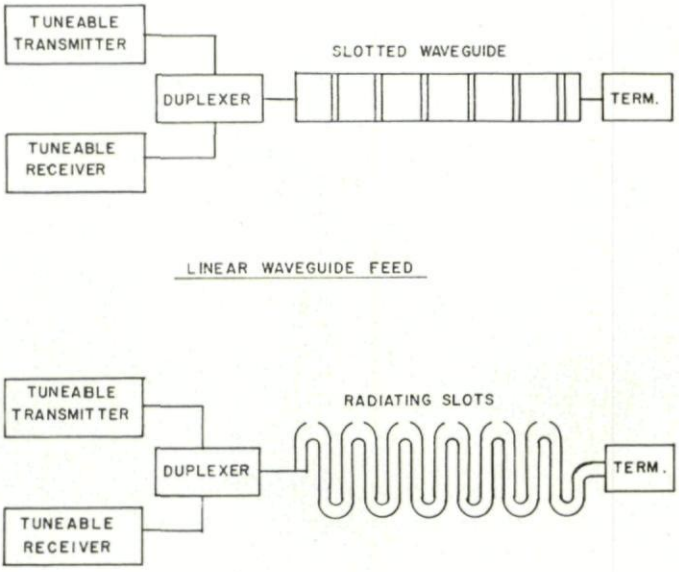


FIG. 2. *Frequency scan linear and serpentine feeds.*

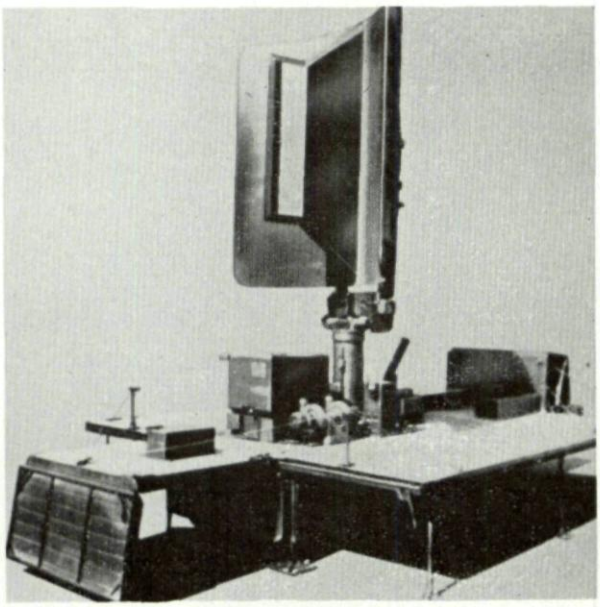


FIG. 3. *AN/MPS-23 radar.*

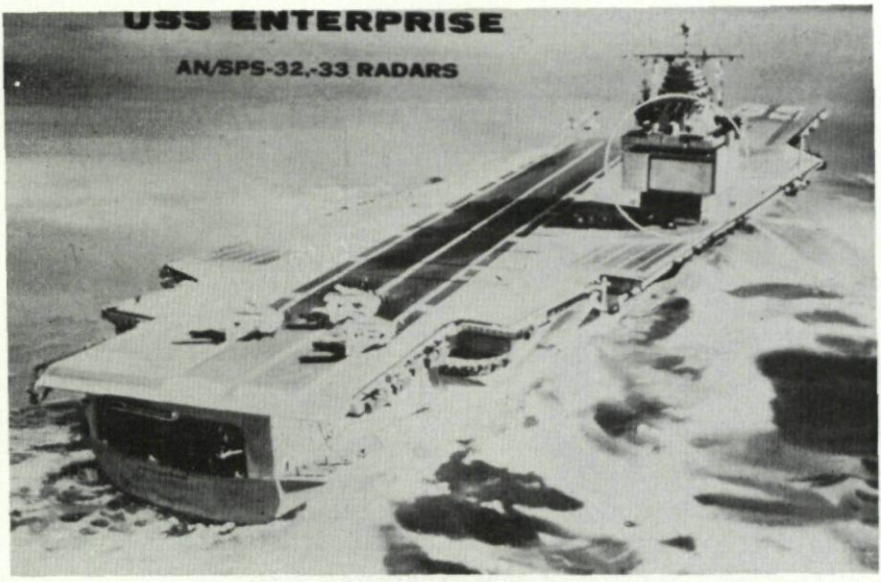


FIG. 4. *USS Enterprise* (*AN/SPS-32, -33 radars*).

which uses frequency scan in one dimension and ferrite phase scan in the other dimension. Figure 5 shows how the two-dimensional phase-frequency scan is implemented in the AN/SPS-33. The input power is fed to a pillbox at whose aperture the energy is picked up by vertical columns of waveguide containing parallel-connected ferrite phase shifters to control the beam

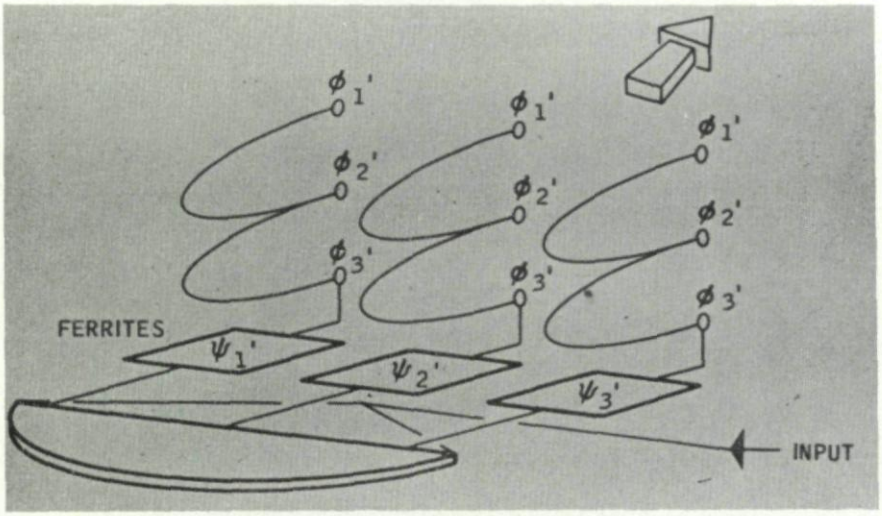


FIG. 5. *Phase-frequency beam steering.*

position in azimuth, and as the frequency is varied serpentine feeds in the vertical columns cause the beam position to scan in elevation.

Figure 6 develops the principle of phase-phase scanning. If a signal is fed through identical series-connected phase shifters there would exist at each tap a cumulative phase shift proportional to the number of phase shift elements through which the signal has travelled. For a two-dimensional scanning system, this portion need only be repeated for the other dimension. The two signals are then combined in a mixer matrix and delivered to the antenna elements through an amplifier chain. By varying the incremental amounts of phase shift the resultant beam may be scanned independently in

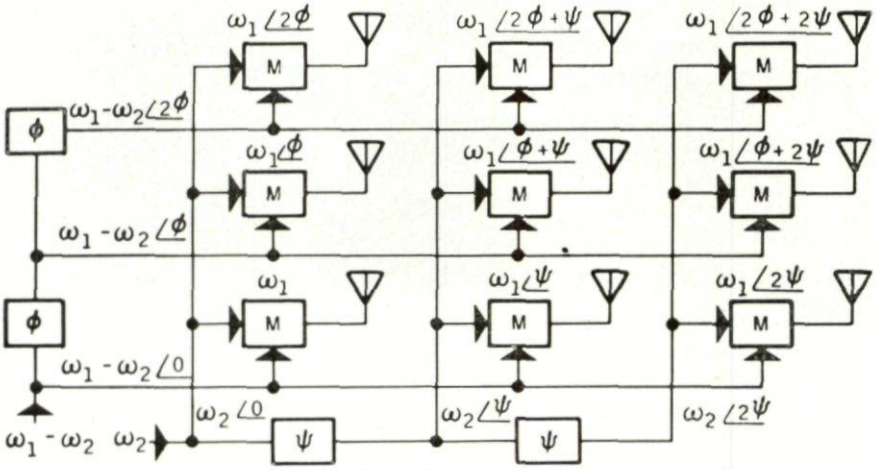


FIG. 6. *Phase-phase beam steering.*

each dimension. Various methods have been investigated for producing the required variable phase shift. One method already mentioned is the series-connected ferrite phase shifter.

Figure 7 shows a linear array of eight ferrite phase shifters. R-f energy enters from the left and travels through each of the phase shifters in series. Phase information is obtained at each of the eight taps. The phase is varied by controlling the magnitude of the d-c current in the coils which are also connected-in series. This means, of course, that all the phase shifters must be nearly identical in phase and amplitude characteristics in order to produce a linear phase gradient across the plane of the array. This requires that rigid quality control of ferrite material be maintained, a rather difficult though not impossible task. The ferrite material must also pass all of the r-f energy which poses some problems when used in a transmit array in terms of power-handling capability of the ferrites. This problem can be mitigated as done in the SPS-33 by the use of parallel-connected phase shifters. Variations in temperature also have an effect on the phase shift characteristics of these devices. Finally, the response time of these devices is relatively longer than other techniques; that is, the time to move the beam

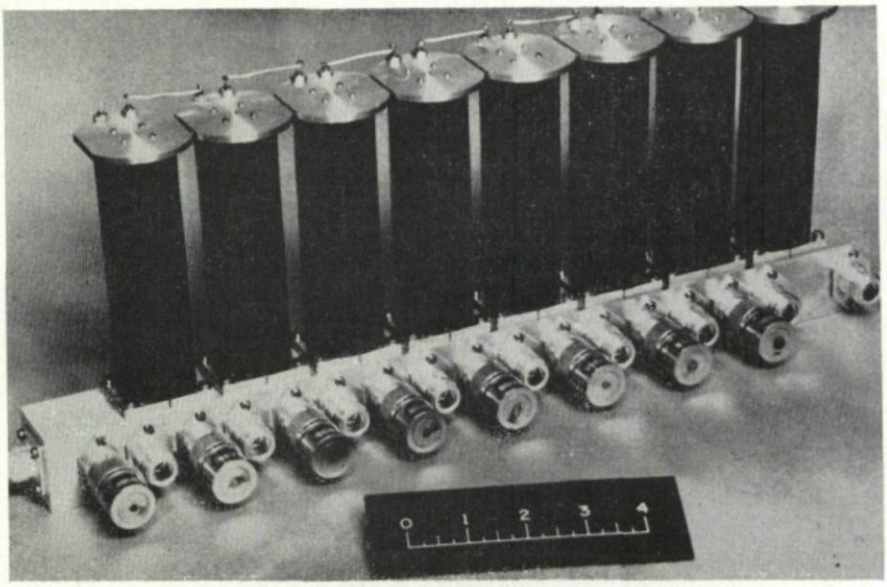


FIG. 7. *Ferrite phase shifters.*

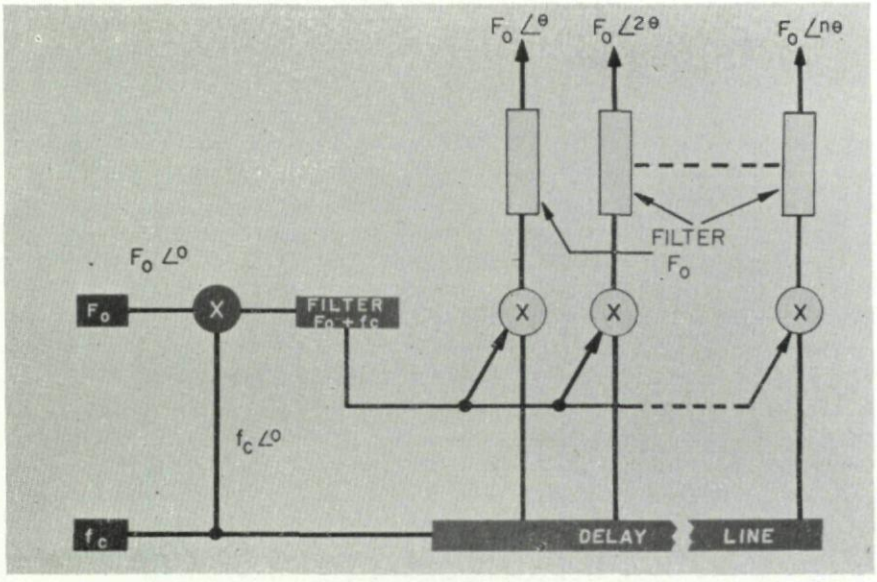


FIG. 8. *Huggins beam steering.*

from one extreme position to another is in the order of milliseconds compared to microseconds when using other phase shift techniques.

Figure 8 illustrates the basic principle of the Huggins-type of phase-phase scanning. The heart of the system is a tapped delay line and a mixer matrix. In the diagram f_c represents a variable frequency and F_0 a fixed frequency. If the variable frequency is propagated down a length of tapped delay line, then at each tap there will be an incremental change in phase whose value is

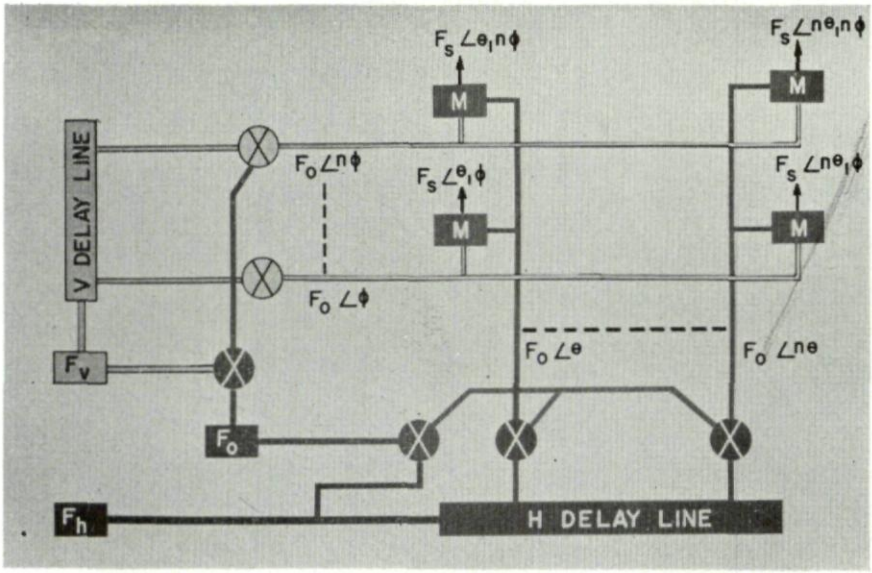


FIG. 9. *Two-dimensional beam steering.*

dependent upon the position of the tap. Since the frequency is being varied, however, the value of the phase at each tap is varying in proportion to the frequency variation. Therefore, if one were to sample the signal at any tap there would exist a variable frequency having associated with it a variable phase angle. If a portion of f_c were now sampled and mixed with the signal from a fixed frequency source F_0 and the resultant filtered, as shown, one would have a sum signal composed of a fixed and variable component. Now, if this sum signal is fed to the mixer matrix and mixed with the tap signals and filtered, we would have at the output of this row of mixers a signal composed of a fixed frequency having a phase angle varying in proportion to the variation of f_c.

Figure 9 shows how this principle is applied to accomplish scanning in two dimensions. The outputs of the vertical and horizontal steering system are applied to a mixer matrix where they are combined before being delivered to the radiating elements.

Figure 10 shows an early implementation of this technique in a 10 by 10 element array. This array generated a ten-degree pencil beam at 500 mc/s.

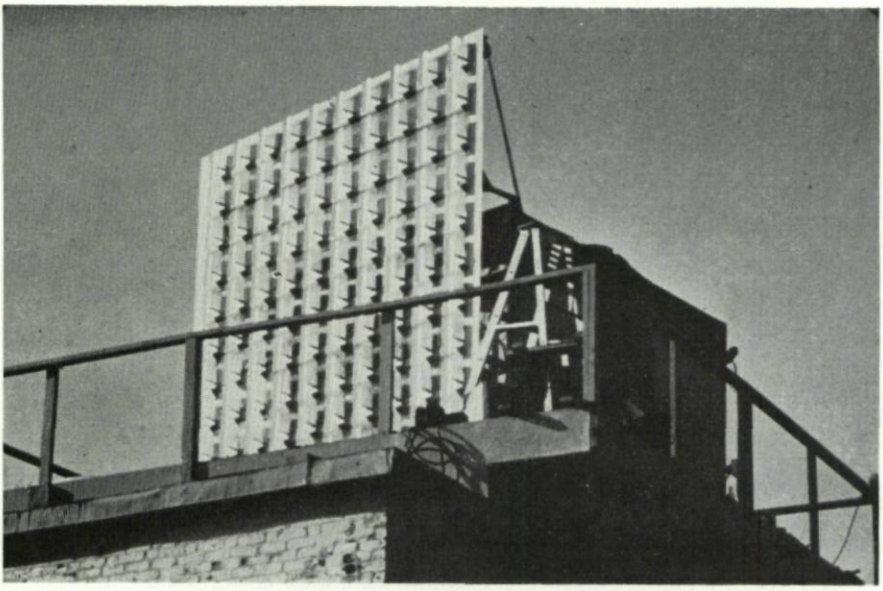


FIG. 10. *Ten-element experimental phased array.*

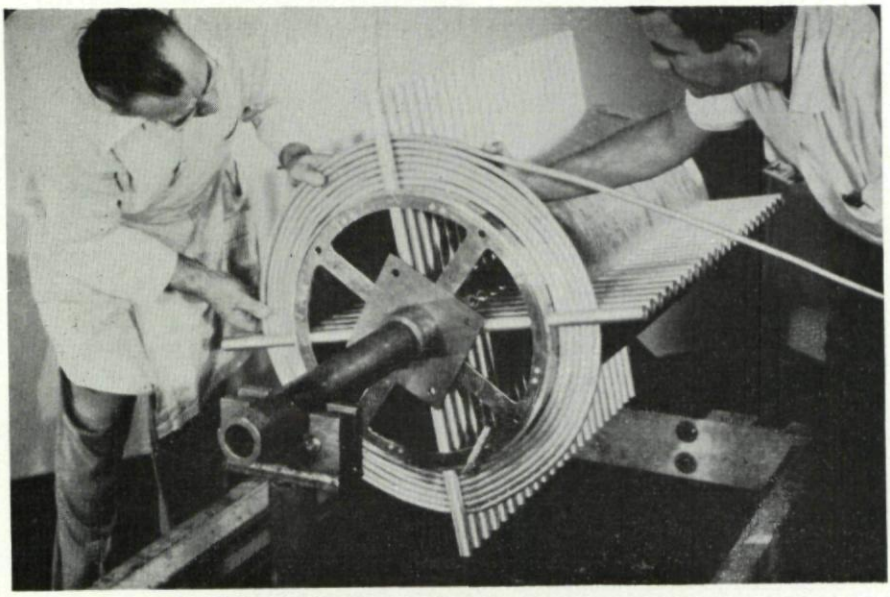


FIG. 11. *Winding ESAR delay line.*

The major objection to this technique is its complexity. The number of mixers required is $2N^2$ where N is the number of elements in a row. The mixers and amplifiers in the system must be capable of tracking both in-phase and amplitude over the band of interest. With careful design, however, these problems have been overcome. Figure 11 shows one form of this delay line being wound. Styroflex cable is being wound in the form of Π sections, one section per tap.

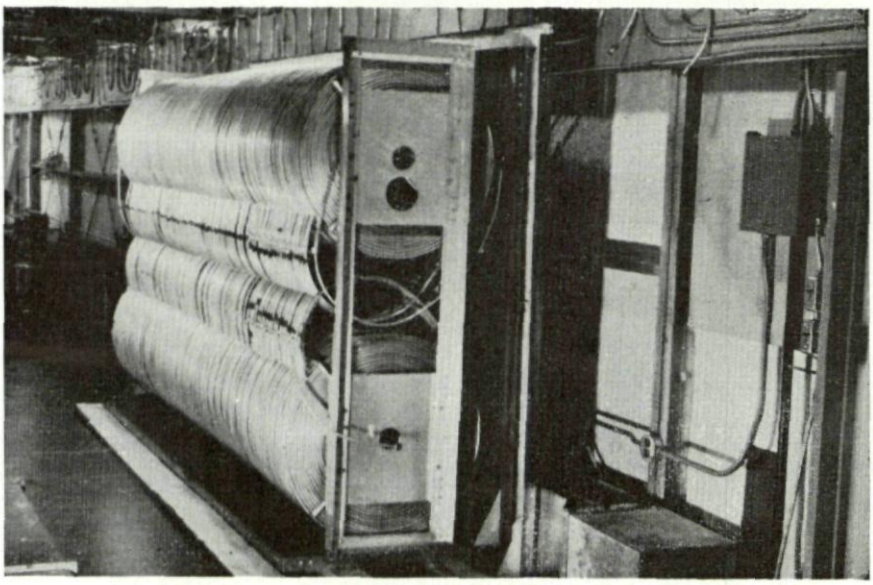


FIG. 12. *ESAR delay line.*

Figure 12 shows a view of the completed line before being enclosed in a temperature controlled cabinet.

Figure 13 shows an outside view of a radar using this steering principle. This radar known as ESAR was developed by the Bendix Corporation under the sponsorship of the Advanced Research Projects Agency and the Rome Air Development Center. Its primary purpose was to demonstrate the feasibility of phased array radars. The radar is completely transistorized except for the transmitter, modulators and other high-power circuitry.

The face of the antenna consists of nearly 8000 elements of the design shown in Fig. 14. This is a log-periodic element designed for L-band operation. These elements are mounted on panels in groups of 64 elements and covered with a foam plastic before being installed in the antenna face. Figure 15 shows a typical panel being prepared for potting. These elements are driven by transmit-receive modules which plug in directly behind the elements. Figure 16 shows a typical module with its cover removed. This package contains the transmitter, driver, steering mixer, duplexer, detector

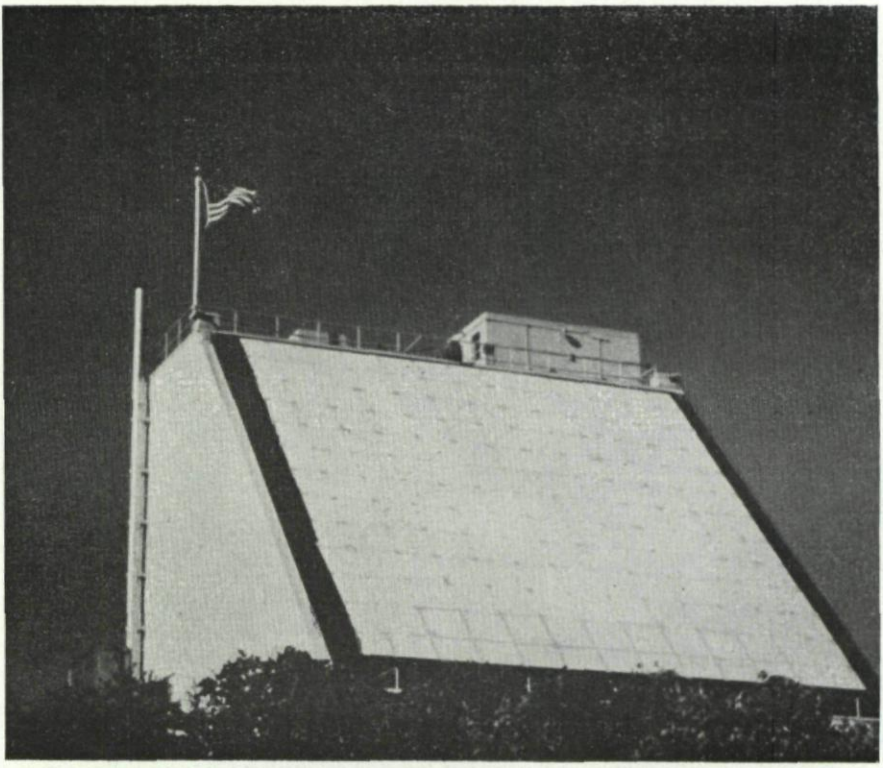


FIG. 13. *ESAR.*

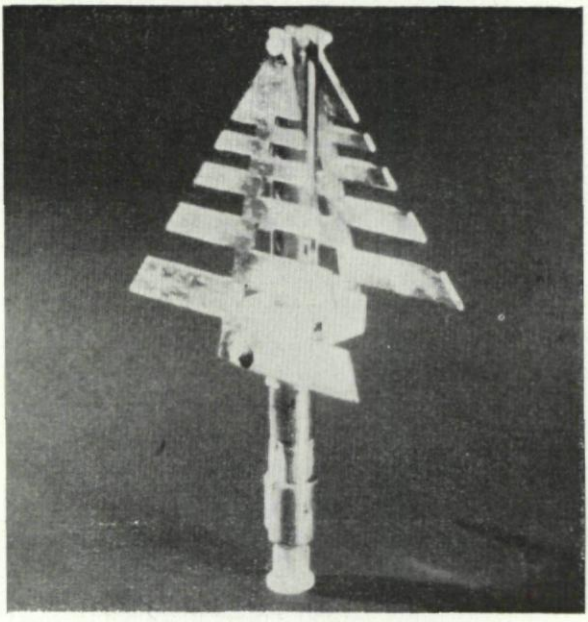


FIG. 14. *Log-periodic element.*

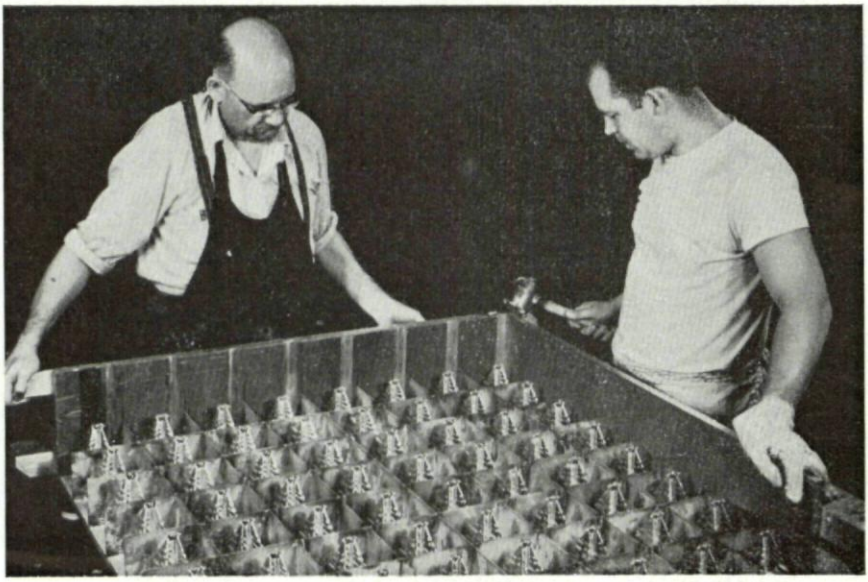


FIG. 15. *Panel of elements.*

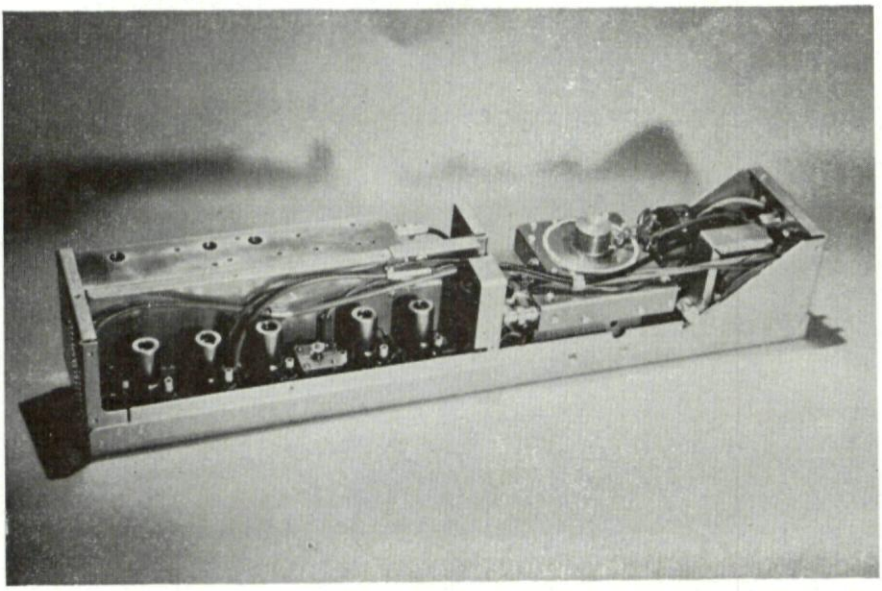


FIG. 16. *T-R module.*

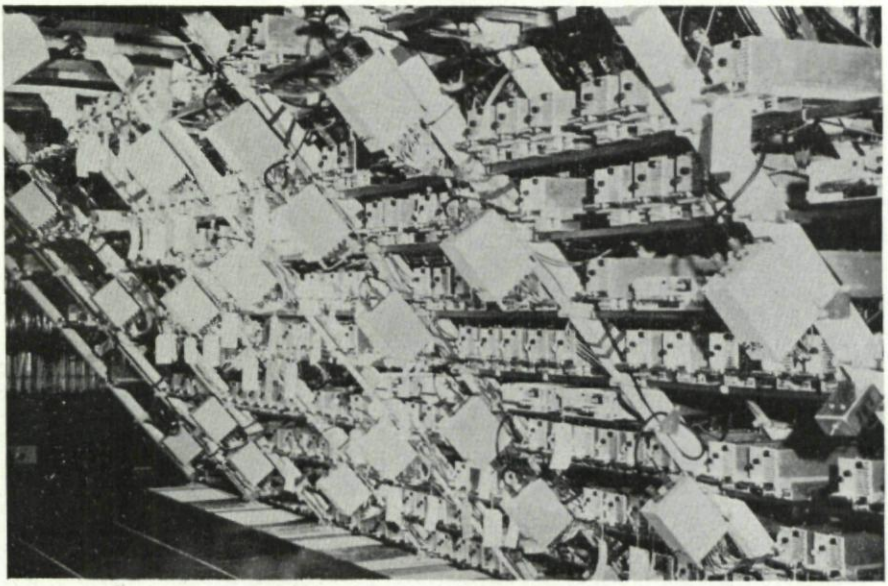


FIG. 17. *Inside view of ESAR.*

and two i-f amplifiers. This radar has only 760 of these modules in use since, being an experimental system and for reasons of economy, there was no need to go to full power. The capability is there, however, if the need arises. The placement of these modules behind the face of the array can be seen in Fig. 17.

Another method for accomplishing electronic beam steering has been under development at the Lincoln Laboratories. This technique involves the use of diodes to switch short sections of delay elements between taps to achieve the required phase gradient. This has the advantage of being more easily adaptable to computer control. This concept, in elementary form, is illustrated in Fig. 18. A number of these sections can be cascaded between taps to give the degree of phase shift required. The diodes are controlled in pairs by computer logic to provide either a direct path or a path through a delay element so that the proper amount of phase shift between taps is achieved. For example, the delay elements can be a short length of cable or stripline.

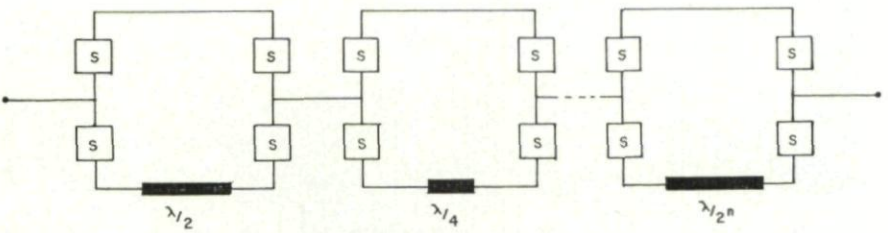


FIG. 18. *Digital phase shifter (schematic).*

Figure 19 shows a view of an experimental 6-bit digital phase shifter operating at 28 mc/s and providing 64 steps of phase shift between 0° and 360° in increments of 5.625° per step. These units would be cascaded between taps. Other configurations can be conceived such as parallel switches and the use of cascaded hybrid sections to which stripline techniques can be applied.

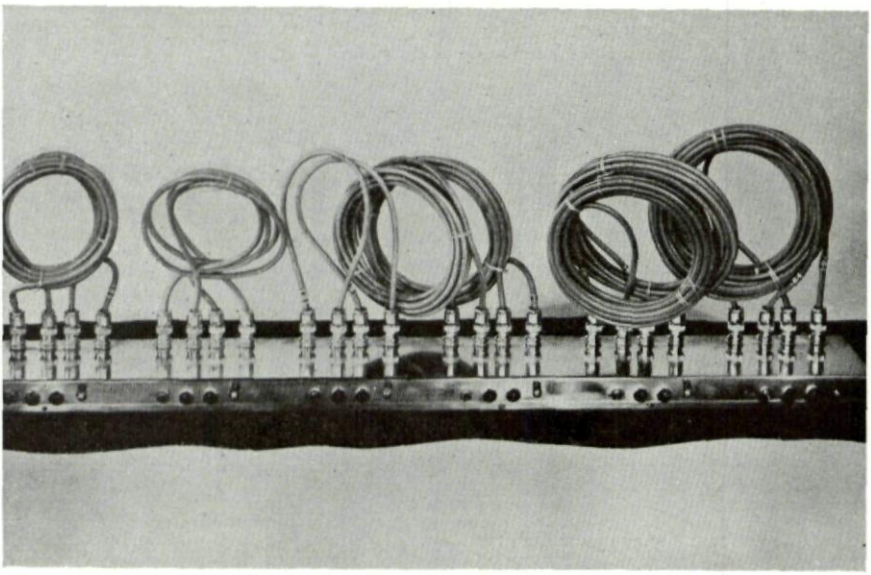


FIG. 19. 28 *mc/s digital phase shifter.*

Figure 20 shows an experimental 220 mc/s phase shifter constructed using stripline techniques.

Another concept of electronic scanning uses the principle of multiple beam formation without the complexity of beam steering circuits. This is based upon the formation of many simultaneous beams fixed in space. A switching matrix is used to select the desired beam output. Angular information from these systems is somewhat less accurate than a steerable array system since the measurement depends upon the target's location in the beam. Perhaps the best known example of this concept is the Luneberg lens shown in Fig. 21. This technique utilizes the dielectric properties of the Luneberg lens to convert a spherical wave front emanating from a feed horn to a plane wave front. By means of a switching matrix, a particular feed horn around the periphery of the lens may be selected to energize a beam in the desired direction. The major problems associated with this technique are those of obtaining uniformity of dielectric material throughout the lens structure and power-handling capability of the dielectric.

Another interesting technique makes use of a delay line at radiated microwave frequency to accomplish multiple beam formation. As shown in

Fig. 22 an array of slotted waveguide lines is coupled to a second array of feed lines by means of directional couplers. Again, by means of a switching matrix, a particular feed line associated with a particular beam direction may be selected. An implementation of this technique is shown in Fig. 23.

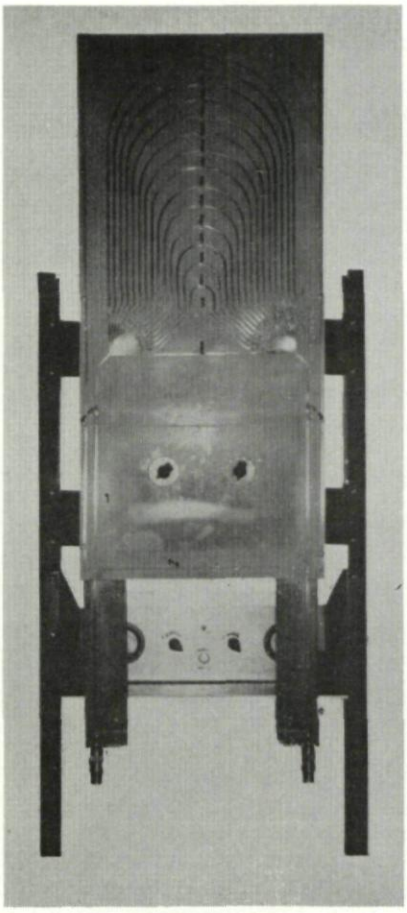


FIG. 20. 220 *mc/s digital phase shifter.*

This is an experimental multibeam height finder known as the ASHR-1. It is a "receive" only system for airport surveillance, depending upon other conventional airport radars to provide target illumination. It operates in the 2700–2900 mc/s region and forms 90 vertically stacked beams. The vertical beamwidths vary from 0.2° for the lowest beam to 1–2° for the highest beam, and each beam has 120° of azimuth coverage. This experimental system was built by Maxson Electronics for the FAA and was based upon an earlier development sponsored by RADC.

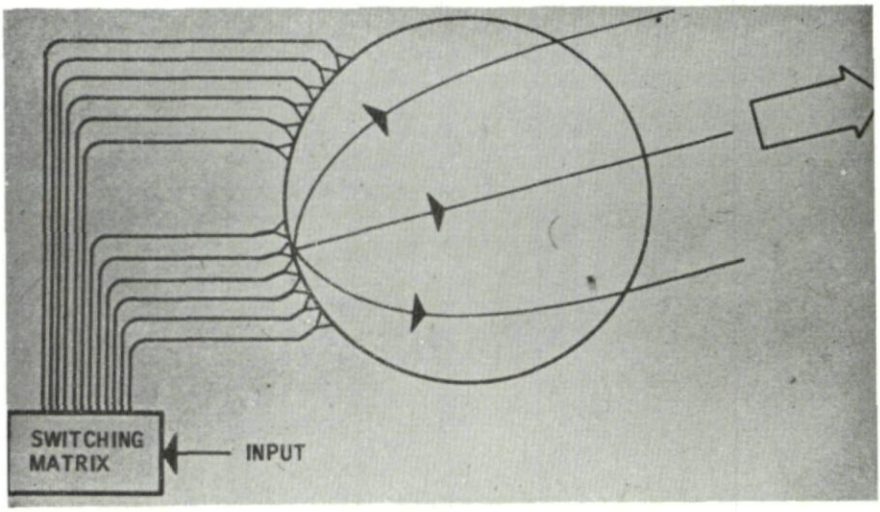


FIG. 21. *Luneberg lens.*

Another form of multiple beam generation makes use of hybrid power dividers. The principle of operation is illustrated in Fig. 24. If an input signal having a magnitude A exists at one port then directly opposite this port there would exist a signal of $\frac{1}{2}$ A in-phase with the input signal, while at the diagonally opposite port the signal would be $\frac{1}{2}$ A 90 degrees out-of-phase with the input signal. Figure 25 illustrates a system for generating four beams. The desired beam is selected by a switching matrix which selects

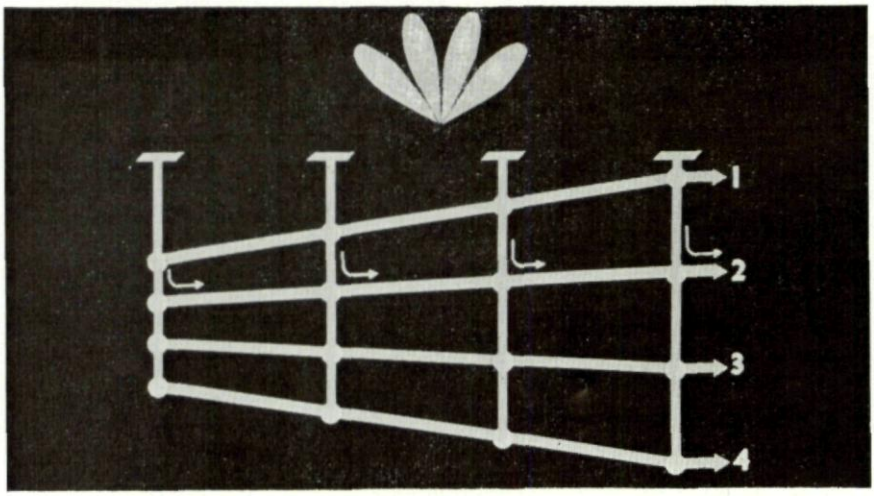


FIG. 22. *Phased line beam switching.*

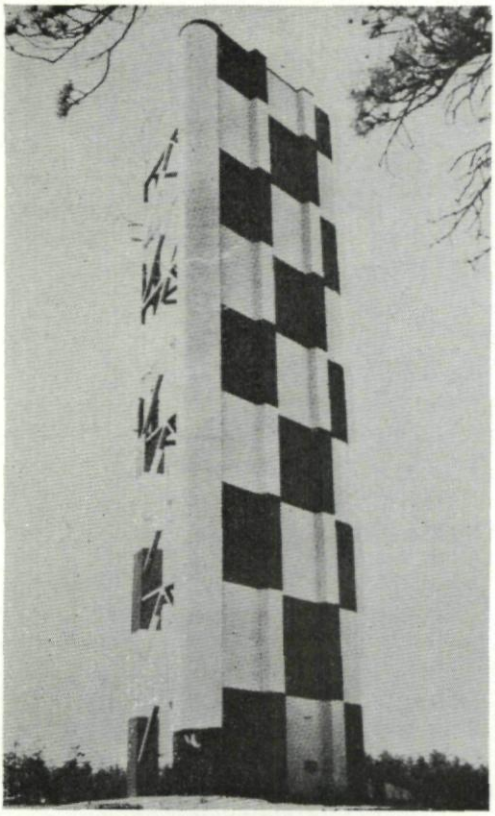
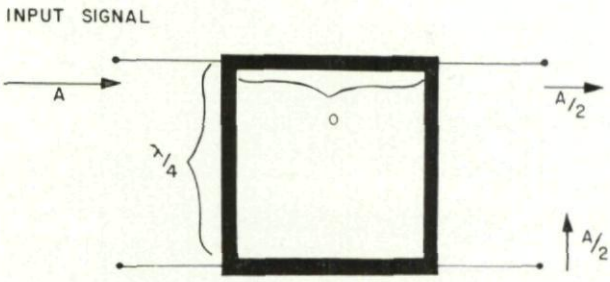


FIG. 23. *ASHR-1 experimental heightfinder.*



HYBRID POWER DIVIDER

FIG. 24. *Hybrid power divider.*

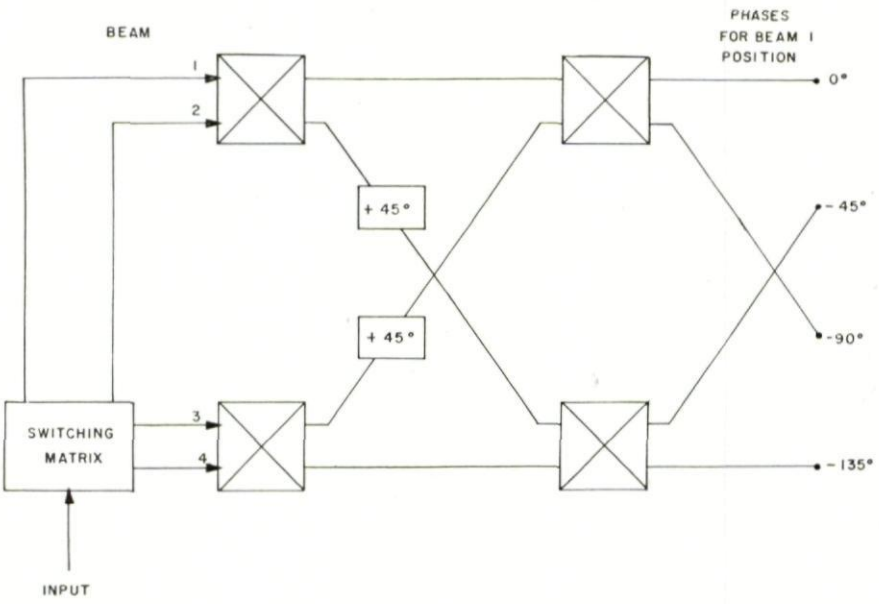


FIG. 25. *Hybrid beam steering.*

the proper beam terminal as indicated. Starting with the uppermost terminal at the left of the diagram, one can trace through the generation of the proper phase front for Beam Number 1 as illustrated on the right of the diagram.

CONCLUSION

Having reviewed the various methods of obtaining electronic scan, it would be in order, next, to examine the question of why so much emphasis has been placed on these developments. Some technical arguments have been presented in the opening paragraphs of this report, but basically, the motivation has stemmed from the recognition by military system planners of the need for systems with growth potential and adaptiveness to cope with the growing complexity of the space age defense threat. A phased array radar properly designed, with growth potential in mind, will permit increases in performance in a more economical and orderly manner than one can expect from a mechanical radar system.

Another factor which deserves comment is the misconception which has arisen regarding the economics and reliability of phased arrays versus mechanical systems. The flexibility and versatility claimed for phased arrays is truly available; however, a phased array system should not be postulated unless performance requirements are such that the capabilities of a phased array are sufficiently exercised. In other works, a phased array should not be considered to do a simple job that can be adequately and more economically performed by a conventional radar. Failure to recognize this fact has too often led to the false conclusion that phased arrays are too expensive.

One must remember that a phased array, properly designed and utilized, is in reality many systems operating from a common aperture, and this multi-function capability should be noted when costing is considered.

DISCUSSION

————: My main interest is in the airborne application of electronic scanning and one of the disadvantages appears to be that one cannot scan to a complete 360° sector. I would ask the authors to give me an indication as to what is the maximum sector that one can cover with electronic scanning and what sort of losses one is likely to get in the aerial going towards the edges of the possible scan.

T. MAGGIO: In a plane array you can scan ±45°, this gives you about a 3 db drop at the edges. Some people have gone ±60° with a greater drop of course. As far as increasing the coverage there is research going on using the Luneburg lens type of array to give you some hemispherical type of coverage. However, there are more problems connected with that. This of course gets to be expensive and requires many many more elements than a planar aerial.

W. HERSCH: Mr. Weiss has gone to considerable length on the pointing accuracy of the beam achievable with a mechanically scanned aerial, can Mr. Maggio give us any indication of the accuracy he can get from an electronically steered aerial?

T. MAGGIO: We have done quite a bit of testing on this point, we have achieved beam pointing accuracies of the same order as conventional radar that is 10th of a beam width, now in the case of the ESAR(MPS 46) it cost several million dollars to find this out. The pointing accuracies were of the order of 0.0050 and for a 1.5° beam width, and tracking accuracies were of the order of 0.01°. These figures represent averages over many months of operation.

REFERENCES

1. ALLEN, J. L., Array radars, *Microwave J.*, **V,** 5, 67–79, 1962.
2. ALLEN, J. L. et al, Phased array radar studies, *Tech. Rpt.* 228 *MIT Lincoln Laboratory*, 1960. ASTIA No. AD-249-470.
3. ALLEN, J. L. et al, Phased array radar studies, *Tech. Rpt.* 236, *MIT Lincoln Laboratory*, 1961. ASTIA No AD-271-724.
4. BLASS, J., The multi-directional antenna: A new approach to stacked beams, *IRE Convention Rec.*, **8,** P. I, 48–50, 1960.
5. SIMPSON, M., New techniques in three dimensional radar, *Trans. IRE*, **MIL-5,** 2, 146–153, 1961.
6. BUTLER, J., and LOWE, R., Beam forming matrix simplifies design of electronically scanned antennas, *Electronics Design*, 170–173, 1961.
7. KSIENSKI, A., Signal processing antennas, *Microwave J.*, **4,** 10 Pt. 1, 77–85, Pt. II, 11, 87–94, 1961.
8. OGG, F., Jr., Steerable array radars, *Trans. IRE*, **MIL-5,** 2, 80–94, 1961.
9. RONDINELLI, L. A., Effects of random errors on the performance of antenna arrays of many elements, *IRE Convention Rec.*, Pt. I, 174–189, 1959.

10. SHARP, E. D., A triangular arrangement of planar array elements that reduce the number needed, *Trans. IRE*, **AP-9**, 2, 126-129, 1961.
11. SHELTON, J., P., Jr. and KELLEHER, K. S., Recent electronic scanning developments, *Proc. of* 1960 *Conference on IRE Military Electronics*, p. 30–36.
12. SHNITKIN, J., Survey of electronically scanned antennas, *Microwave J.*, Pt. I, **3,** 12, 67–72, 1960; Pt. II, **4**, 1, 57–64, 1961.
13. AULOCK, W. H. VON, Properties of phased arrays, *Proc. Inst. Radio Engrs, N.Y.*, **48,** 10, 1715–1727, 1960.

CHAPTER 22

ELECTRONICALLY SCANNED ANTENNA SYSTEMS, II.

E. A. KILLICK
Admiralty Surface Weapon Establishment
and
D. E. N. DAVIES
University of Birmingham, England

A general account is given of the basic methods of electronic scanning of aerial beams. The feed complexity of a large phased array can be reduced by the grouping of elements which are fed together. Other simplifications result from space tapering and array thinning. The phase relationships in the array may be varied by the use of continuously variable or quantised phase shifters. Frequency scanning can achieve the same result. An account is given of various techniques for the formation of multiple beams.

1. INTRODUCTION

The paper by Diab and Maggio describes the main systems of electronic scanning or steering of radar beams and illustrates the advanced engineering of such systems in the U.S. The aim of this paper is to add some general comments about electronic scanning systems and their limitations without reference to particular examples.

We take electronic scanning to mean any electronic means of moving an aerial beam within microseconds or at most milliseconds (and will in general assume that the same beam is used both for transmit and receive). Electronic scanning does not appear likely to be attractive merely to replace mechanical controls in simple radar systems; its main application appears to be in comparatively sophisticated and expensive radar systems with some degree of computer control. The capability of a radar system which exploits the agility of an electronically controlled beam may, in some applications, be much greater than that of a system using the same transmitter power but limited to mechanical rates of beam positioning; a favourite example of this is the ability to combine in one radar the functions of long range surveillance with high data-rate tracking of selected targets.

The current problems of electronic scanning appear to be:

(1) *Cost* (above all): particularly of the potentially most versatile aerial systems scanning a 1° or 2° beam over wide angles in two planes and using large numbers (typically $\sim 10^3$ or 10^4) of aerial elements and of phase shifters or amplifiers. This paper will not attempt to comment on estimated costs or the possible mass-production of components. "Array-thinning,"

"grouping of elements" and "quantised-phased" are techniques which will be mentioned which can somewhat reduce the cost and the control complexity of large phased arrays.

(2) *Restrictions on Instantaneous Bandwidth:* To form a pulse of length t seconds requires a spectrum of at least $1/t$ cycles per second and this of course applies equally if the pulse is formed by pulse-compression techniques from a modulated transmission of much longer duration. An aerial system which has adequate instantaneous bandwidth receives the components of this spectrum back from the target in their correct relative amplitudes and phases. To see what will happen one can inspect the two way pattern, gain and phase of the aerial system for each C.W. component of the modulation system to be used and make the appropriate summations.

The widest instantaneous bandwidth obtains in aerial systems where the beam is collimated and steered by true time delays; for instance an unstepped lens or mirror ideally only shows slow changes in gain, beamwidth and scan over a wide spectrum. However, in phased arrays which do not employ multiple amplifiers it is only possible because of phase shifter loss to insert changes of up to 2π radians or one wavelength delay. The relative delay required between the ends of an array whose beam is scanned $\theta°$ away from the normal is $\sin\theta \times$ the length of the array and may be typically up to $\sim$50 wavelengths; in that case phase shift settings in unit of 2π will be quite in error at the ends of the array if the frequency is changed by $\pm 1\%$. This difficulty is likely to become of concern when the pulse spectrum is $\sim$100 mc/s ($t = \frac{1}{100}$ μs).

Evidently severe restrictions on instantaneous bandwidth are inherent in frequency-scanning systems unless large frequency changes are used to produce the scan. Typically an aerial may be designed to scan an angle equal to its main beamwidth for a change in frequency of 1 mc/s, and deterioration in received pulse shape and apparent aerial gain will occur when the pulse spectrum is about 2 mc/s ($t \sim \frac{1}{2}$ μs).

(These effects can alternatively be described and analysed by considering the transit time of the pulse along the array instead of the pulse spectrum.)

(3) *Resistive Loss and Peak Power Limitations:* both these restrict the use of phase shifters and switches, and loss is a problem in the long line lengths of frequency scanners and of some phasing and switching networks. Typical estimates of resistive loss in various scanning systems are in the range 1–3 db one way. This loss increases the aerial noise temperature as well as decreasing aerial gain. In switches and phase shifters of a suitable design to use in fairly large numbers ($\sim 10^2$) peak power limits are typically in the range 10–50 kW. These problems are of course by-passed in systems using multiple amplifiers both on transmit and receive.

2. PHASED ARRAYS

The two simple arrangements of phase controls and radiators are shown in an elementary one dimensional manner in Figs. 1 and 2. If the phase controls are put in series as in Fig. 1 they must be very accurate and stable

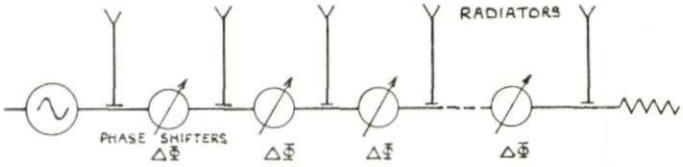


FIG. 1. *Series arrangement of phase shifters.*

since errors may be cumulative; loss in phase shifters in series would normally make amplifiers necessary at each radiator. Phase controls are more likely to be used in parallel as in Fig. 2 and only phase shifts of up to 2π radians used unless amplifiers are added: some such parallel feed is visualised in speaking of "quantised-phase" and "grouping of elements." An alternative parallel feed as in Fig. 3 is likely to be attractive at centimetric wavelengths. A two dimensional version might look akin to a waveguide lens with a phase shifter in each cell: a one dimensional version would be contained between parallel plates. A monopulse (static split) system can be combined with the electronic scan by using a four or two-horn feed as in a conventional lens or paraboloid aerial.

The phase controls may be applied at intermediate frequencies with a mixer and amplifier at each radiator. Some systems of this sort are mentioned under Receiving Arrays: similar transmitting arrays can, of course, be designed.

To avoid secondary beams (grating lobes) and achieve optimum gain the planar array surface should consist ideally of phase controlled radiators spaced apart not further than $\lambda_u/1 + |\sin \theta|$ where λ_u is the free space wavelength and θ is the maximum angle to the normal to which it is wished to scan. (If the array surface is curved the question of spacing and polarisation of the radiators is more complex.) To give a $2° \times 2°$ beam widely scanned

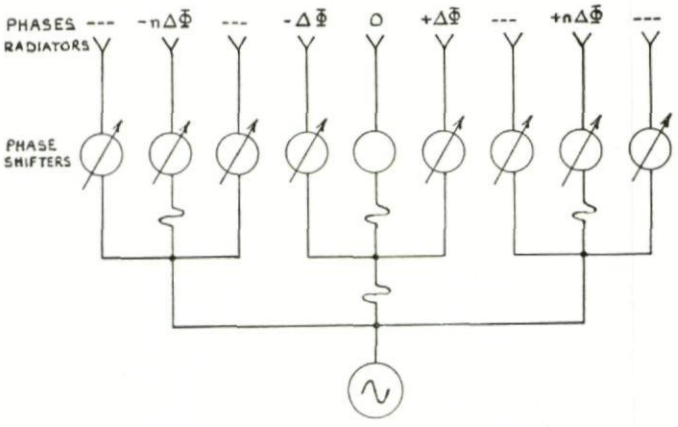


FIG. 2. *Parallel arrangement of phase shifters.*

in both planes the planar array will consist ideally of a grid of approximately 70×70 phase controlled radiators spaced little more than $\lambda_u/2$ apart. The radiators may be any of the familiar wide-beam radiators such as dipoles over a ground plane or open-ended waveguides, or end fire elements such as log-periodics. The mutual coupling properties of these will appear to be somewhat similar when spaced about $\lambda/2$ apart: the mutual coupling

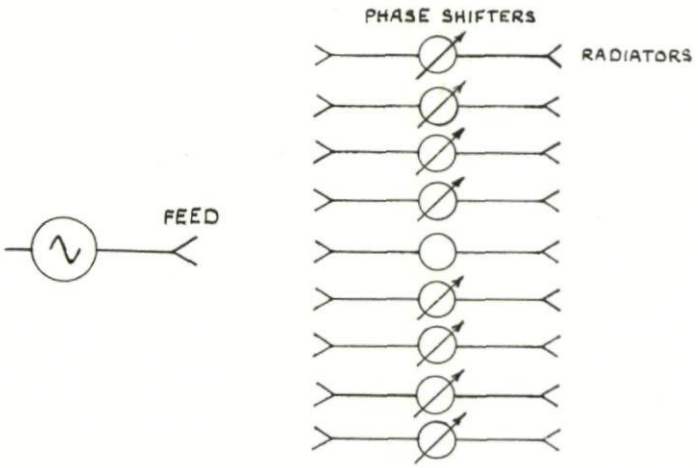


FIG. 3. *Parallel arrangement of phase shifters.*

(probably around 20 db) does not appear to be an outstanding problem in the design of large arrays. If the $2° \times 2°$ beam is to be scanned in one plane only, the array will consist of approximately 70 phase controlled radiators each of which will probably be a simple linear array (such as a slotted waveguide) in order to form the beam in the unscanned plane. Scanning in two planes involves the major problem of sheer numbers and density of phase control elements, but it is not likely that the power to be handled by each control element will be large: typically not greater than 1–2 kW peak except for multimegawatt systems. Scanning in one plane only may sometimes mean that the power to be handled by each control element (possibly up to ~50–100 kW peak) becomes the outstanding problem. The problems of designing phase control elements are eased if amplifiers, even in small numbers are introduced into the system (together, of course, with circulators or separate arrays for transmit and receive).

2.1 *Grouping of Elements*

If the radiators of the array are taken in groups, and if each group is phase-controlled from one control element the main array will appear when scanned to consist of a number of large radiators whose spacing is greater than the requisite $\lambda_u/1 + |\sin\theta|$ and secondary beams will appear. However each group of elements has itself a rather narrow beam because of its

size and this reduces the amplitude of the secondary beam whilst also reducing the gain of the scanned main beams: a compromise can be reached for a certain range of scan by tailoring the group beam shape and possibly also by scanning it. The groups may also be interconnected. The attraction of sub-dividing the array design in this way is that it becomes possible to use a limited number of amplifiers and main phase controls: for instance the main phase control to each group could be a true time delay of many wavelengths of lossy line followed by an amplifier, after which low loss 2π phase shifters could be used to steer the group pattern.

2.2 *Space-Taper and Array-Thinning*

A narrow beam aerial is usually designed with a field amplitude taper across it to reduce the sidelobes. In an array it is possible instead to reduce the density of radiators at the array edges whilst keeping the amplitude at each constant, and hence economise on radiators and elements by some 30% or 40% in a square array. The increase in beamwidth is similar to that due to the normal amplitude taper thought the sidelobe level achieved may be somewhat worse. The spacing of the "thinned out" radiators must be randomised. Random "thinning" can also be extended over the whole array in a manner not contributing to space taper and the beamwidth will not then increase. However, the power that would have appeared in secondary lobes, had the missing radiators not been randomised, appears in low level wide angle sidelobes and the array gain falls. Taking amplitude tapers into account, the gain of an array is always proportional to the number of active radiators present until their spacing approaches half a wavelength.

2.3 *Quantised Phase*

Some simple and attractive forms of phase control give switched increments of line length or phase shift rather than continuous control: for instance, a "five bit" phase shifter gives on-off increments of 180°, 90°, 45° $22\frac{1}{2}$° and $11\frac{1}{4}$° and can be set to give phase shifts up to 360° in steps of $11\frac{1}{4}$°. Much coarser control than this may be acceptable at the elements of a large array (possibly "three bits" or 45° steps) particularly if some arrangement is made to avoid any probability of phase jumps appearing at repetitive intervals across the array.

3. PHASE SHIFTERS

Two main types of phase shifter are being developed:

Phase shifters of one main type consist of ferrite-loaded waveguide, coaxial or strip-line with a magnetic control field. The properties of ferrite materials (including in this term any low loss magnetic material) vary rather strongly with frequency since they depend upon the gyromagnetic resonance in the microwave band and upon a domain wall resonance in the 100 mc/s region. Ferrite phase shifters appear most suitable for frequencies about 3000 mc/s, although they can be made for much lower frequencies. Ferrite materials show a complex and frequency dependent form of loss (due

to the growth of spin-waves) when the microwave magnetic field exceeds some critical value and this rather strictly limits the maximum possible peak power density in the material; the maximum possible mean power depends on cooling and is a lesser problem. All ferrite materials are to some extent temperature sensitive and some degree of temperature control or compensation is necessary. Unless the magnetic energy due to the control field and resulting flux in the ferrite is kept low, which means low control fields and small volume, the power required from the control circuits during fast resetting of many phase controls will become excessive. Design compromises have to be made between the demands of low control fields and small volume and those of low microwave loss and sufficient peak power handling.

A microwave ferrite phase shifter being developed is the "latching" phase shifter. This uses rings of ferrite within a waveguide or helix. The rings can be switched in less than a microsecond between the left-handed and right-handed states of remanent magnetisation. Current pulses of one or two amps are typically required for switching and no "holding" currents are required since the ferrite rings have high remanent magnetisation ($\sim$80% of saturation magnetisation) and zero demagnetisation factor. As each ring is switched, ferrite phase increments are produced which can be arranged as described under "Quantised Phase." The latching phase shifters are normally nonreciprocal devices (since the sign of the magnetisation does not control phase shift in a reciprocal device) and the switching currents must be reversed between transmit and receive. A small "four-bit" ($22\frac{1}{2}^\circ$ steps) phase shifter of this type (which might be used in such an arrangement as Fig. 3) might typically be designed at a frequency of about 6000 mc/s with a loss of 1 db, a peak-power limit of a few kilowatts, and an average control power consumption of a few watts. A reciprocal phase shifter which does not demand transmit/receive switching normally requires a "holding" current from the control (since the magnitude and not the sign of the magnetisation controls the phase shift). To provide a large pulse during the rise time of the magnetic field and a low voltage "holding" current thereafter is a problem in the design of fast control circuits and the control power consumption may be some tens of watts if the circuit is to be able to re-set the phase in microseconds.

The design of both types of ferrite phase-shifter can be "stretched" with increase of size, control power, and possibly loss to handle peak powers up to around 50–100 kW.

Phase shifters of the second main type consist of lengths of delay line "shorted" by voltage controlled diode switches in some such arrangement as is shown in Fig. 4. Phase shifters using PIN junction diodes as switches appear most suitable for frequencies below about 6000–3000 mc/s although they can be used for higher frequencies. The cheaper PN junction diodes may also be usable in the lower frequency microwave band. The diodes must have low capacity and resistance, and a high ratio of forward to backward impedance; for high power operation these demands must be combined with that of a high breakdown voltage (greater than the peak to peak signal voltage when the diode is back-biassed). The diodes appear to be best

designed for low impedance (5–10 ω) lines. A typical four-bit PIN diode phase shifter might be designed at a frequency of about 3000 mc/s with a loss of 1–1.5 db and a switching time of much less than a microsecond. At present the upper limit of peak power handling in the microwave band

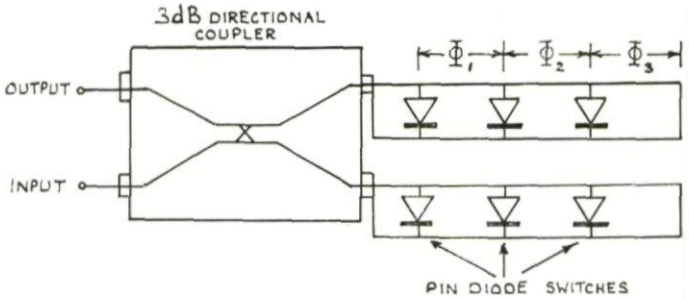


FIG. 4. *Switched phase shifter using pin diodes.*

appears to be ~10 kW (unless many diodes are used, each contributing small amounts of reactance to the line). Junction diodes require an adequate heat-sink but no temperature control.

4. FREQUENCY SCANNING

The two restrictions in the design of frequency scanning arrays are the loss in the delay line and the restricted instantaneous bandwidth or minimum pulse length (discussed in the introduction). The latter restriction can be put simply: the aerial however designed must not scan much more than a beamwidth for a frequency change equal to half the spectrum of the pulse, the exact limit depending on how well the pulse spectrum and hence the range resolution and the apparent aerial gain are to be preserved. Systems can of course be devised in which the transmission is frequency modulated during the pulse sufficiently to produce a scan of many beamwidths (such a scan-within-the-pulse is mentioned under "Receiving Arrays") but in that case very much less than the whole pulse spectrum is being used for range resolution.

The length of delay line L required between successive radiators of a frequency scanning array as shown in Fig. 5 is given by:

$$\frac{\dfrac{L}{\lambda_g}}{\dfrac{d}{\lambda_0}} = \frac{\sin\theta}{\dfrac{\Delta f}{f}}$$

where d is the radiator spacing and θ is the angle from the normal to which it is desired to scan (at frequency f) by a frequency change Δf: λg and λ_0 are the delay line and free space wavelengths respectively at frequency f. This

relation, which is exact only if the velocity in the line is constant, indicates that the significant property of the line is its loss per wavelength and this is minimised by using air filled line (probably waveguide) of the largest possible cross-section. With careful design folded waveguide or helically wound waveguide can be made with a loss approaching the theoretical

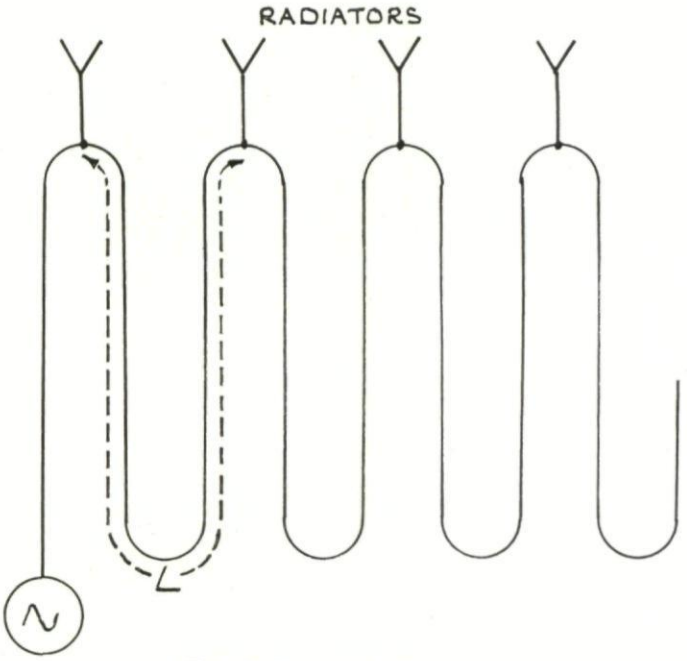


FIG. 5. *Frequency scanning array.*

value for the same length of straight line. The resistive loss in the array is approximately half that in the uncoupled line and is evidently approximately proportional to

$$\frac{\text{scan in beamwidths}}{\text{percentage frequency change}}.$$

Typically an S band waveguide array with a 2° beam scanning 45° for a 5° frequency change may have a loss of about 1 db. If such an array were scaled to another band the loss would be expected to vary as $\sqrt{f}$.

L can be made very short by using a line in which the velocity varies rapidly with frequency but this appears likely always to be more lossy than sufficient length of line with low dispersion.

Since a frequency scanner is sensitive to a fractional change in wavelength it is equally so to a fractional change in dimension and is therefore temperature sensitive. Typically a frequency scanner must be temperature controlled to within a few degrees C or else the temperatures must be monitored and corrections computed to the angle/frequency calibration.

5. MULTIPLE BEAM-FORMING TECHNIQUES

A high gain receiving aperture will receive energy only from a small range of incident angles (neglecting the energy received on side lobes). Energy incident on such an aperture from other directions will be re-radiated or absorbed in parts of the aerial circuitry. However, by suitable design of the aerial system it is possible to set up many spatially overlapping high gain directional beams from a single aperture so that energy may be received from (or transmitted from) separate outputs corresponding to different directions. These outputs may be simply compared, or they may be combined in various ways to give sum beams at intermediate directions or having various beam shapes.

The most obvious method of producing multiple beams is by using a lens or mirror collimator with multiple scanned feeds. Lenses of constant or slightly varying refractive index, such as waveguide lenses, can be designed to scan to many beamwidths with some deterioration of pattern and gain. Lenses of continuously varying refractive index can be designed to scan perfectly any angle: for instance the Luneberg lens is a sphere whose refractive index is a function of radius and brings an incident plane wave to a focus on the opposite surface of the sphere, hence multiple feeds over part of the surface can produce multiple identical beams over a large solid angle. Luneberg lenses have been built using foamed dielectric of varying density and using artificial dielectrics such as aluminium particles dispersed in foamed dielectric. These lenses appear suitable for receiving, but moderate power levels may cause overheating and destruction of the centre of the lens. The "tin-hat" or parallel plate geodesic lens is a robust analogue of the varying refractive index lens which collimates, in one plane only, by virtue of the varying path lengths over a curved parallel plate surface.

Reflectors can be designed to scan many beamwidths with some deterioration in pattern and gain if the focal length is sufficiently long, though the design of efficient feeds becomes difficult. Alternatively a spherical reflector can be designed to produce identical beams over a large solid angle but with rather low aperture efficiency.

Multiple beam forming systems using a lens or reflector collimate and scan by true time delay and may have very wide instantaneous bandwidth, though account must be taken of the fact that at different frequencies the beamwidths and directions of the beams will change, keeping their relative cross-over points unaltered. Certain lens systems may however have restricted bandwidth due either to the necessity of zone stepping the faces or to restricted bandwidth performance of some artificial dielectrics.

Multiple beams may equally well be formed from linear or planar arrays by appropriate interconnection of phase shifts or delay lines between elements. Two examples of such networks are shown in Figs. 6 and 7. Such networks do not in general provide true time delays.

It can be shown that in order for close spaced multiple beams (originating from a single aperture and each with the full aperture gain) to remain independent their directional patterns must be a set of spatially orthogonal

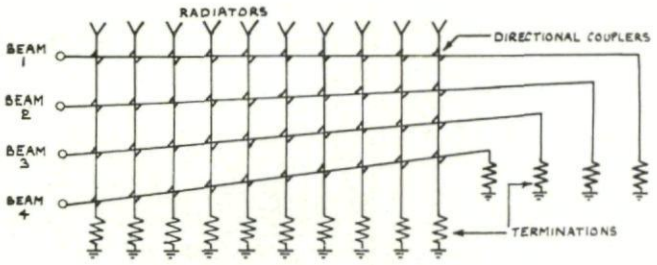


FIG. 6. *Series fed beam forming matrix.*

functions. This applies equally whether the beams are formed by multiple feeds to a lens or mirror or by an array network. For instance a group of $\sin x/x$ patterns (i.e. the pattern radiated by an untapered amplitude distribution) are spatially orthogonal if the patterns are displaced from each other by slightly more than a beamwidth, having cross-over points at 4 db below the peaks of the beams. If it is attempted to make the cross-over points of the beams higher than the level set by the orthogonality criterion for the pattern there will either be coupling between the beams or the gain of each beam will drop (for instance due to spillover in the multiple feed design or to loss in the directional coupler loads of a network). The orthogonal spacing of $\sin x/x$ beams corresponds to a total phase shift of π radians along the array between adjacent beams. This gives the result that the maximum number of independent beams is equal to the number of elements

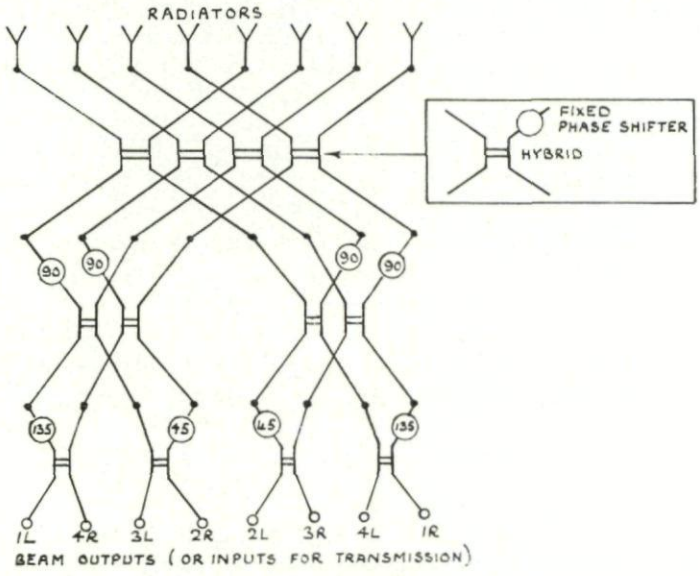


FIG. 7. *Parallel fed beam forming matrix (Butler).*

in the array provided that the element spacing is not reduced below $\lambda_0/2$. The effect of any tapering is to reduce the number of beams available from a given array. The theoretical gain of each independent beam of a multiple beam array is given by the full gain of the aperture in that particular direction. In other words the only loss is that due to the projected view of the aperture away from broadside.

6. ELECTRONICALLY SCANNED ANTENNA SYSTEMS

Although an orthogonal set of sin x/x curves represent the best least-square fit of the target distribution this is not always the sole requirement particularly when in the presence of large extended targets. Therefore in many cases it is desirable to employ amplitude tapers, to reduce side lobe levels, there will in any case always be some form of taper on a lens or reflector aerial. The series multiple beam network of Fig. 6 may be designed with an amplitude taper by suitable choice of the coupling values of the directional couplers. The parallel network (Butler matrix) provides sin x/x directional patterns without tapers. However, one simple method of achieving the tapering on the Butler matrix is to feed two adjacent beams in parallel, this gives the equivalent result to a cosine amplitude taper across the array. Feeding three adjacent beams is equivalent to a cosine squared taper. Clearly this method of tapering may be used on any multiple beam aerial. It would also appear that the best results for two way multiple beam formation would involve different directional patterns on transmission and reception, thus involving non-reciprocal devices in the beam forming matrix.

6.1 *Switching between Beams*

Fig. 8 shows a possible switching network to select one beam from a group of $2N$ beams; $2N - 1$ switches are required, and N switches are in series to

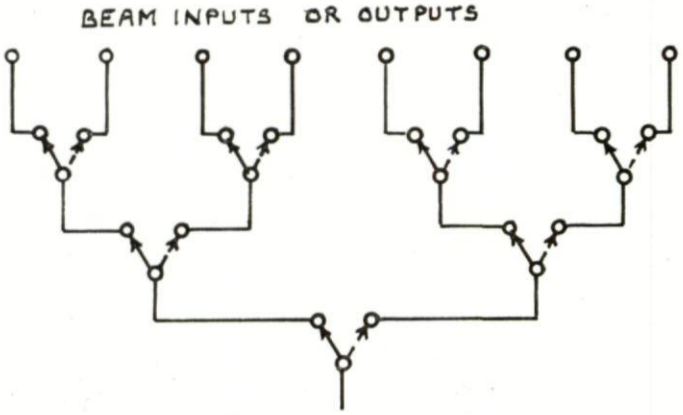


FIG. 8. *Switching matrix.*

any one beam input. Very similar considerations apply to the design of these switches as those discussed under the heading of "Phase Shifters:" suitable fast and easily controlled ferrite or junction diode switches with a few tenths of a db insertion loss can be designed to carry peak powers up to some tens of kilowatts. Some of the most convenient ferrite switches are non-reciprocal. Ferrite switches can be designed to carry peak powers of megawatts if the control fields used are large ($\sim$1000 gauss): to reset many such switches in microseconds or milliseconds would require considerable control power.

7. ELECTRONIC BEAM-SCANNING FOR RECEIVING ARRAYS

It is evident from the section on multiple beam-forming networks that a large number of independent beams may be formed from a single aperture and that this process may be adopted for both transmission and reception or for reception alone. In the latter case it is often necessary to deploy one receiver for each network output corresponding to an independent beam direction. For search purposes the output of these receivers would be examined sequentially and this can be accomplished without loss of information only by the use of very high sampling rates as stipulated by the sampling theorem. Thus considering a transmitted pulse length of t and a receiver bandwidth of $1/t$, the information relating to target position will be given by a bandwidth of $1/2t$ (after idealised demodulation) so that each receiver would have to be sampled sequentially at a rate of $1/t$ in order to retain all the information at a single output. If there are N receivers the duration of sampling at each will be t/N and a bandwidth of N/t will be required in the sampling circuit. However the signal-to-noise ratio is established before sampling by the receiver bandwidth of $1/t$. It is also possible to produce the same effect by the use of a single fast scanning receiving beam which samples all the beam directions in space at a sampling rate (or scanning rate) of $1/t$. This technique is usually known as within-pulse scanning and has been used very successfully in the sonar field. The advantage of the system is that the circuits required for beam deflection operate only for reception at low power. The volume of space to be searched is illuminated by a wide beamwidth floodlight transmission which results in the overall directional characteristic of the radar being given by the one way directional pattern of the receiving array. The receiving beam scans repetitively the entire sector (or solid angle of search) in the direction of the transmitter pulse. It can be shown that the signal-to-noise ratio and consequent range performance of a within-pulse scanning radar may be exactly the same as a conventional slow scan radar with the same power transmitter and aerial aperture, provided that idealised coherent integration is employed in both cases. However the within-pulse scanning array receives the same mean power in the form of a larger number of smaller energy pulses and its overall performance is therefore critically dependent upon the efficiency of integration process. The integration process becomes an outstanding difficulty if many beamwidths scan are

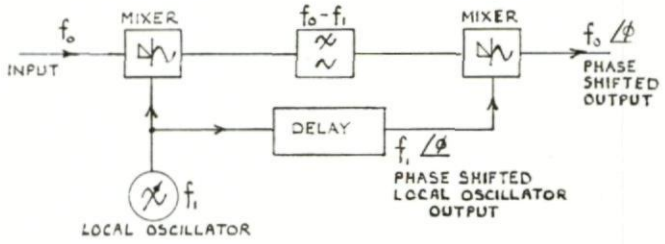


FIG. 9. *Transferred phase shift system.*

required. For non-fading targets and incoherent integration there may be a loss of performance but for fading targets and high probabilities of detection this loss may be very small.

It will be evident from the above that this form of scanning is a surveillance system. It is not possible to vary the power radiated at any given target as with a beam-positioning system. The application of such a technique appears to be relevant to uniform search for initial detection or to provide height information within fan beam cover of a mechanically rotated radar.

Several techniques are available for achieving the high scanning rates necessary for within-pulse scanning (scanning rates in the region of 2 kc/s to 10 mc/s). One technique known as the transferred phase shift system and sometimes known as the Huggins phase shifter makes use of a variable frequency local oscillator and a delay line as shown in Fig. 9. This may be adapted to beam steering the directional pattern of a linear array by the arrangement of Fig. 10. It can be seen that on reception there is no need to remove the frequency modulation before demodulation unless the bandwidth is so large that circuit techniques become too difficult. In order that the signal-to-noise ratio of the scanning receiver is not worsened by the noise

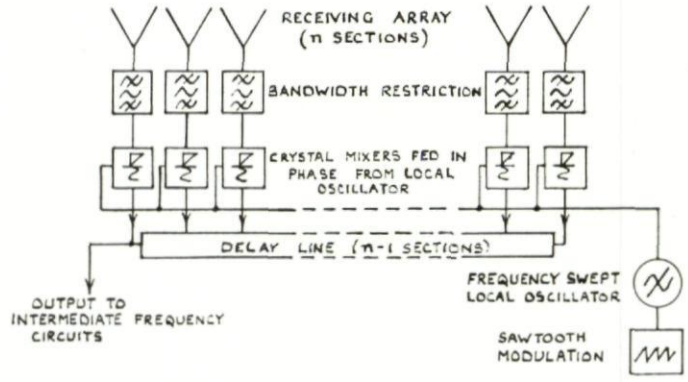


FIG. 10. *System for electronic scanning or tracking of a receiving beam.*

in the wide bandwidth circuits it is necessary to fix the signal-to-noise ratio at each element of the array by means of amplifiers and/or filters, so that the signal-to-noise ratio is not degraded by the subsequent wide bandwidth mixers.

It is also possible to use frequency scanning for within-pulse scanning (this corresponds to transferred phase shift with a variable frequency in the medium). In such a case the transmission must consist of a frequency modulated transmission which illuminates the region of interest by scanning the transmitting beam across the region very fast. In this case the effective transmitted pulse length in each direction is given by the time that the transmitter was illuminating that direction. Some interesting features of this technique are that the same array may conveniently be used both for transmission and reception and it is also convenient to combine a "chirp" type pulse compression with this form of scanning.

An interesting alternative method of producing fast repetitive scanning is by successive frequency translation of the output of the elements of a linear array. This produces continuous repetitive scanning since the array factor component of the directional pattern moves uniformly along the sin θ scale. Figure 11 shows a schematic diagram of such an arrangement employing single sideband modulators to effect the frequency translation.

When any form of the transferred phase shift system or frequency scanning is employed the total frequency excursion Δf used to produce the phase change must be several times larger than the scanning rate $1/t$ multiplied by the number of elements N in the aerial array:

$$f \gg N \cdot \frac{1}{t}$$

This is because if Δf is too small the delay line required to produce the phase shift is too long and the transit time along it exceeds the sampling time t/N. It is interesting to consider the effect of employing variable true time delays (rather than fixed time delays and variable frequency) for scan-within-the-pulse. Continuous fast scanning would correspond to a continuous high rate of change of delay in the delay lines which would mean that short input pulses to such a lines would suffer pulse lengthening or compression and furthermore the pulse from each element would suffer a different amount of this distortion. It would therefore appear that for very short pulse applications it would be better to use fixed beams formed by the time delays with post detector beam sampling switches.

A technique known as aperture modulation can also be used to produce either multiple or scanning beams or to produce effective amplitude tapers. One example of such modulation is to switch on elements of a linear array sequentially one at a time. The resultant radiated pattern consists of a family of orthogonal sin x/x patterns with each pattern at a different radiated frequency. The frequency difference between adjacent radiated sin x/x patterns is given by the switching frequency. If each element is excited at a slightly different frequency the far field pattern is one continuously scanning

beam and this can be considered as the exact Fourier transform of the arrangement of Fig. 11. If the aperture consists of one moving element (or several with sequential sampling) then the far field pattern consists of multiple beams at different frequencies. Low sidelobes may be obtained by switching on and off the elements in such a fashion that the resultant time average power distribution along the array corresponds to the required taper distribution. The reduction of sidelobes by this method may be applied in addition to conventional tapers to produce very low sidelobes.

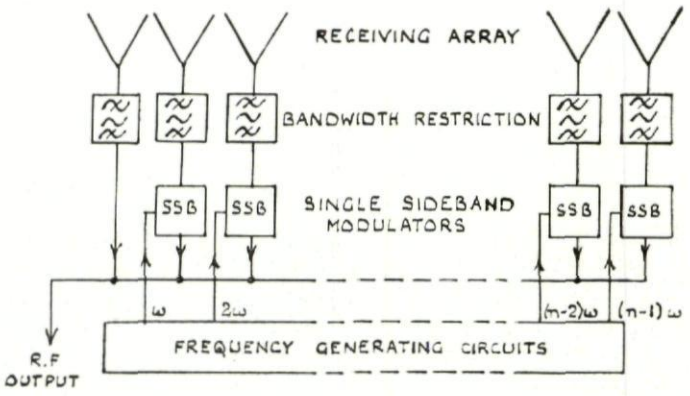


FIG. 11. *System for continuous electronic scanning of a receiving beam.*

Although within-pulse scanning has been described in terms of a floodlight transmission with a scanning receiving beam, the roles could be reversed though clearly the former system has the advantage of simple circuitry for the high power point of the equipment.

Within-pulse scanning can be applied to crossed arrays. If two linear arrays are crossed at right angles and their outputs are multiplied together, the multiplied output consists of a pencil beam with the same main beamwidth in both planes as that produced by a rectangular filled aperture with sides equal in length to the two linear arrays. This beam can be scanned in both planes by phase control of the $2N$ elements of the crossed arrays (in comparison with the N^2 elements of the rectangular aperture. However the gain of the crossed array is that due to the area of the two linear arrays and is therefore much less than that of the rectangular aperture. Since the multiplication process applies only to received signals this technique is applicable to a within-pulse scanning receiving array for use with floodlight transmission. Problems can arise with such a system when two or more targets are present at the same range but different bearings or elevations; in such situations multiplication cross products can occur indicating the presence of false targets but the extent of these products is dependent upon the statistical properties of the targets.

CHAPTER 23

HIGH POWER C.W. RADAR TRANSMITTER

WALTER S. BAUMGARTNER
Jet Propulsion Laboratory, Pasadena, California, U.S.A.

This report is principally a description of an operational high-power CW radar transmitter. To properly orient this transmitter with the NASA-JPL space program, some preliminary discussion of the space telecommunications problem, the existing Deep Space Network, and in particular the continuing program of advanced system development is presented. The system in which the transmitter functions is described and its use as a planetary radar and as a proving ground for equipment is discussed. The design parameters for such a transmitter are delineated and the history of its development is traced. Particular attention is given to problems encountered and to their solutions. There is a detailed description of the hardware involved. The present plans for increasing the radiated power still further and upgrading the transmitter system in general are presented.

1. INTRODUCTION

The Jet Propulsion Laboratory (JPL) of the California Institute of Technology is assigned by the National Aeronautics and Space Administration (NASA) responsibility for certain unmanned lunar and planetary space programs. In support of these programs, a Deep Space Instrumentation Facility (DSIF) is operated with tracking and data acquisition stations located at approximately 120 degree spacings around the Earth. These stations provide continuous radio and radar contact with space vehicles. The present locations of these stations are Woomera and Canberra, Australia; Johannesburg, South Africa; Goldstone, California, USA; and a new installation scheduled for Madrid, Spain. The purpose of these stations is to provide tracking of spacecraft (two angles, Doppler and range), issue commands to spacecraft, and receive telemetry and television signals. Figure 1 is an aerial view of the Echo Site at Goldstone, California and is representative of the other stations. The 85-ft HA Dec antenna operates with a 10-kW *S*-band transmitter and a low-noise receiver.

II ADVANCED SYSTEM DEVELOPMENT

To develop improved systems and subsystems for these stations, and to obtain important data on the planets of our solar system, a special advanced development station is operated at the Venus site, Goldstone, California.

Figure 2 is an aerial view of this site. Two antennas are available. The 30-ft Az-El antenna in the foreground is presently used at X-band and higher frequencies, and the 85-foot Az-el antenna is used with the high-power transmitter for lunar and planetary work. Near this antenna are the buildings housing the hydraulics for the servo motors, the 1.5 MW heat exchanger and pumps, and the transmitter power supply and controls. The main control

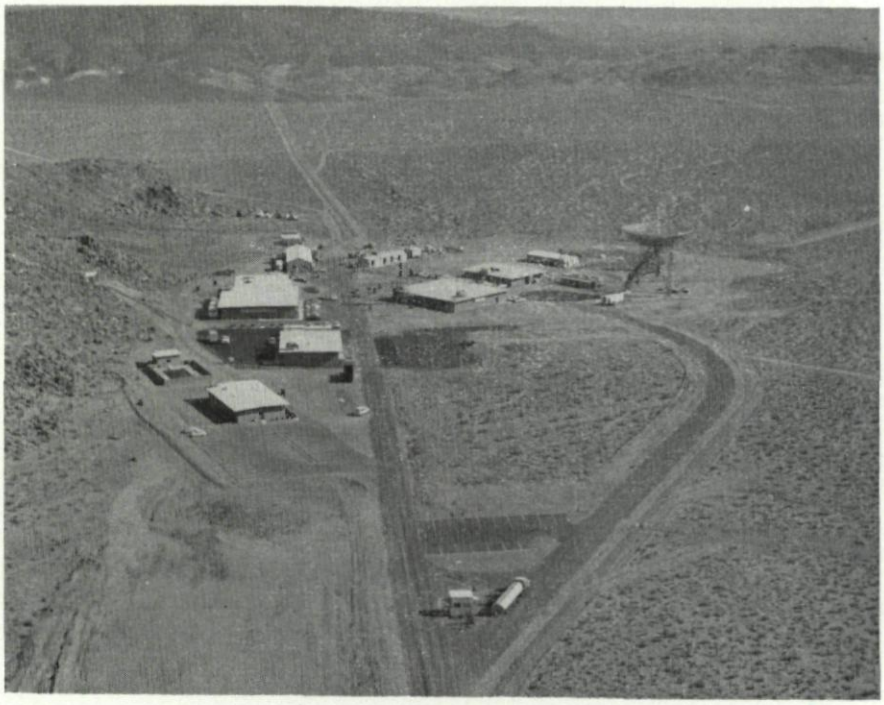


FIG. 1. *Aerial view of Echo site.*

building is in the foreground, 1500 ft from the antenna to avoid radiation hazard. Figure 3 is a view of the interior of this control building showing a portion of the high-power radar installation. There are also limited office and laboratory facilities at this base and a permanent staff providing 3-shift, 24hr, 7-day operation. Engineering development work is accomplished at Pasadena, California. At any time, several engineers from that facility may be at the Venus site. Firm power at 60 c/s is furnished by a line from a commercial power system. As the pictures indicate, this area is desert (Mojave) and all water is supplied via tank truck. Ambient temperatures can range from $+16°$F in winter to $+130°$F in summer. High winds are common. The site offers the advantages of adequate land, good isolation from other installations, and low radio noise.

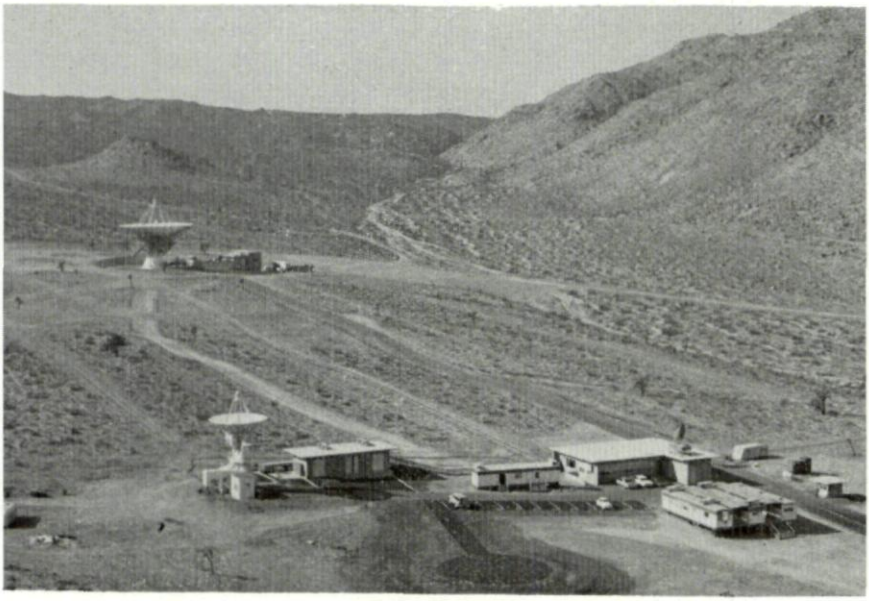


FIG. 2. *Aerial view of Venus site.*

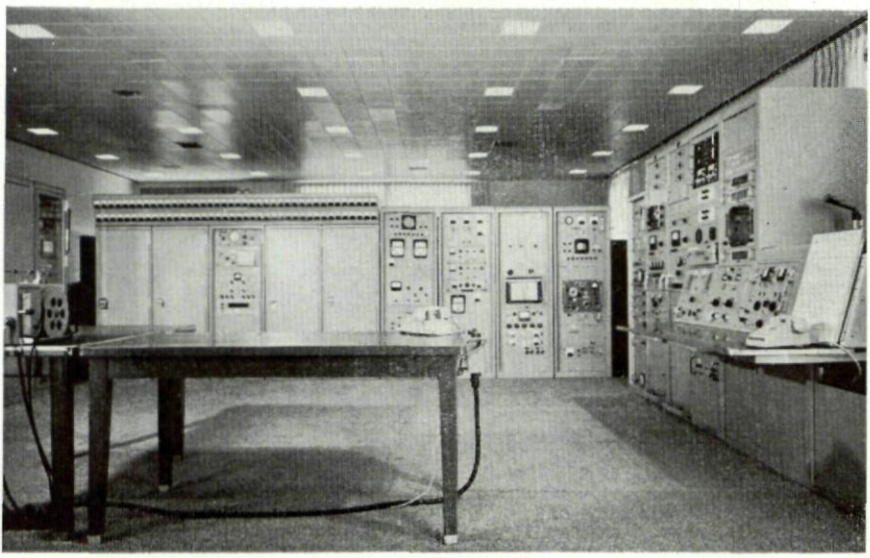


FIG. 3. *Planetary radar control room.*

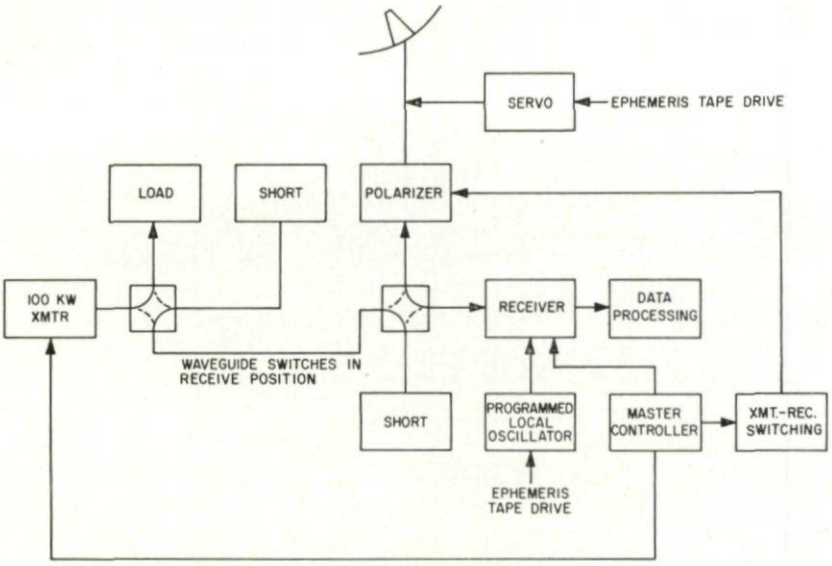


FIG. 4. *Planetary radar block diagram.*

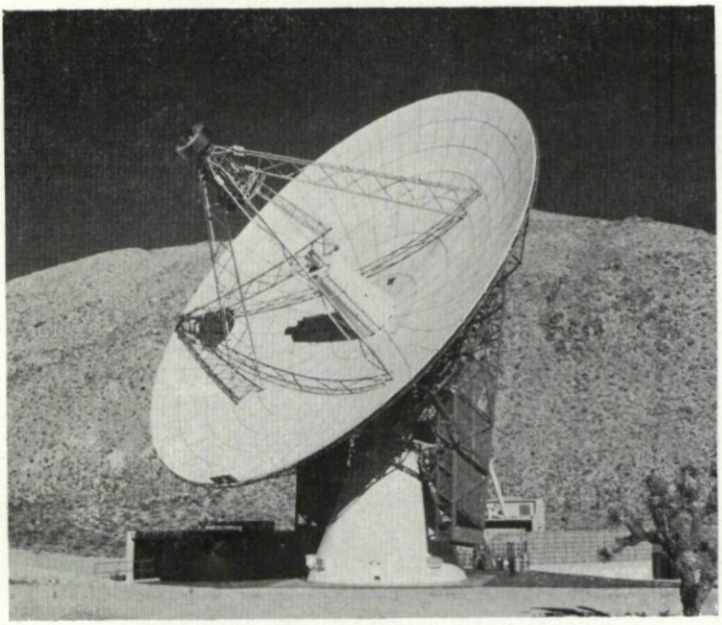


FIG. 5. 85-*ft Az-El antenna.*

As mentioned previously, this site serves the dual functions of field-testing of advanced systems and accumulation of astronomical data. The first function is met by continuous operation (10 to 20 hr per day, 7 days a week) and the building of considerable flexibility into the equipment. The planetary radar work has resulted in refinement of the Astronomical Unit and correction in ephemeris information which has contributed to the success of space vehicle missions.

The major components of the high-power CW planetary radar are shown in block diagram form in Fig. 4. It is apparent that the system is a monostatic radar using the same antenna for transmission and reception. In planetary work, the round trip signal time is from 4 min to over 1 hr so the change from a transmit to a receive configuration may be accomplished by waveguide switches without the switching time becoming an appreciable part of a cycle. In addition to the operation of the waveguide switches the

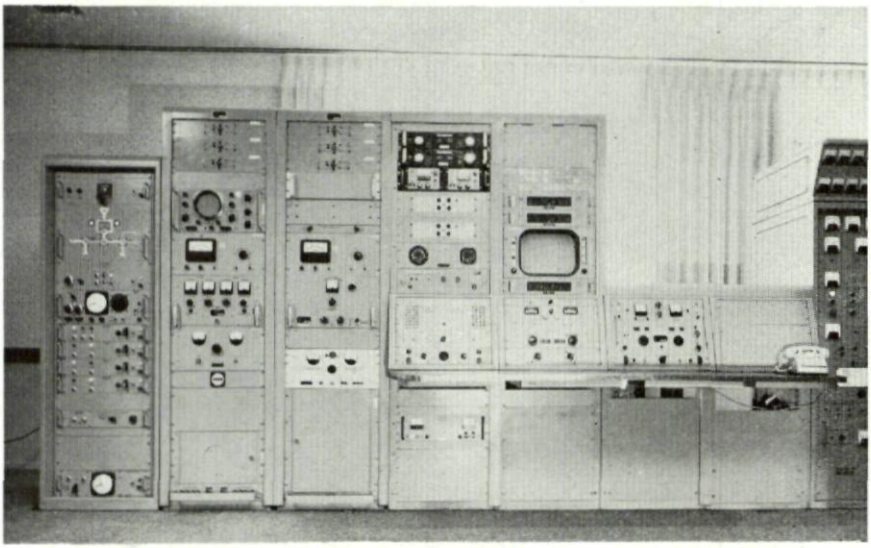


FIG. 6. *Antenna servo console.*

receiver local oscillator is off during transmission and the transmitter drive is off during reception.

Starting at the top center of Fig. 4, the antenna is an 85-ft Az-El steerable parabola which, at the operating frequency of 2388 mc/s provides 54-db gain and a beam width of 0.35 deg. The antenna and its Cassegrain feed are shown in Fig. 5. The small object at the apex of the quadripod is a 6-ft tunneled antenna used in a bistatic configuration on some lunar experiments.

The servo system of the antenna is driven by error signals developed from an ephemeris tape prepared in advance by JPL for the planets of interest and the time of operation. The punched paper tape is converted by digital

equipment to azimuth and elevation signals which drive the hydraulic servo motors. A more recent development uses an on-site computer to drive the antenna and eliminates the need for prepared tapes. The servo control and boresight TV console are shown in Fig. 6.

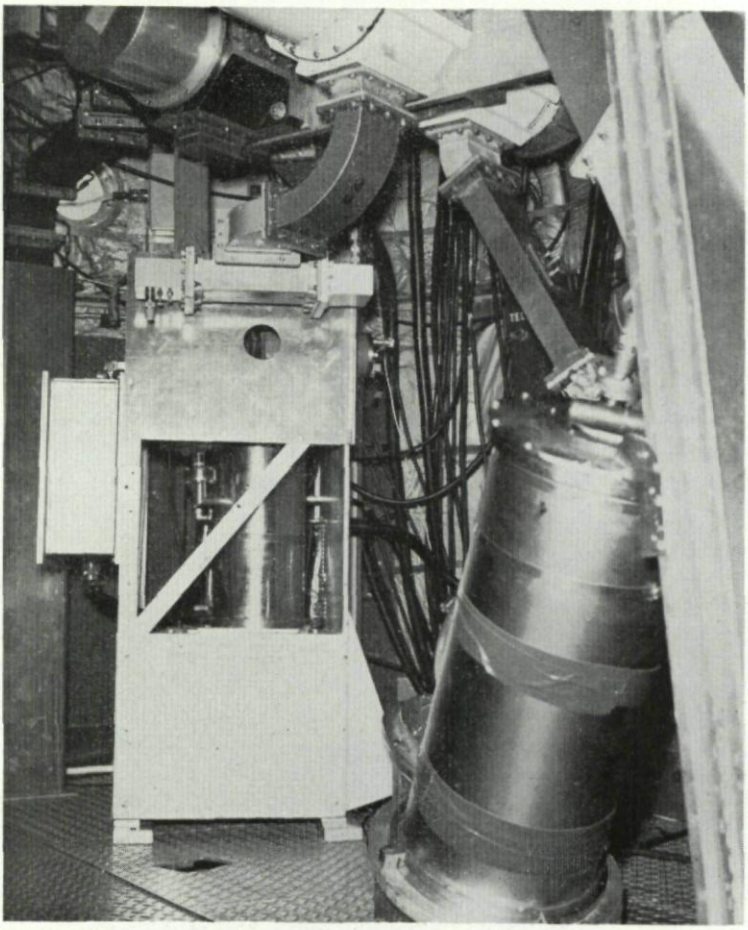


FIG. 7. *Cassegrain feed cone installations.*

3. THE RECEIVER

The receiver used in this system is a low-noise triple conversion superheterodyne with five channels capable of handling various types of modulation. The first amplifier, a traveling-wave liquid helium-cooled maser, is located in the Cassegrain feed cone on the antenna (shown in Fig. 7). It

provides a system noise temperature at zenith on cold sky of 29°K. With this noise temperature, the receiver has been operated open loop to −190 dbm and is capable of maintaining phase lock in a 5-c/s loop to −176 dbm. In addition to the maser, the receiver first converter, RF switches and polarizer, and a signal generator are located in the cone. The signal is converted to 30 mc/s and conducted to the control building via coaxial cable. Figure 8 is a view of the portion of the receiver in the control room. Figure 9 shows

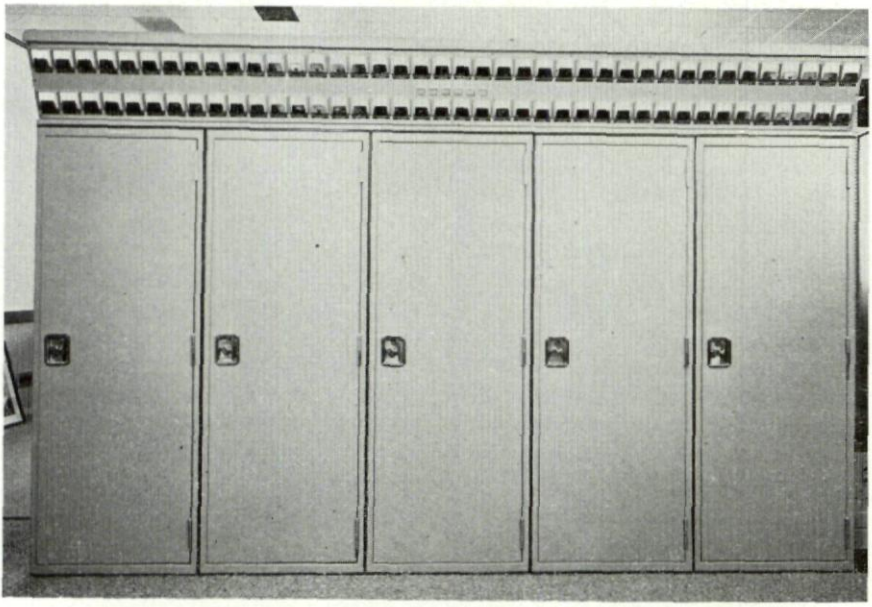


FIG. 8. *Receiver Mod IV.*

the type of modular construction used in the receiver, the exciter, and other radar components. Each of the gold plated modules shown comprises some complete circuit element such as a phase detector, mixer, amplifier, etc. All have 50-Ω input and output and are shielded and filtered to maintain leakage levels below 1 μV around the box and on the power leads. The gold plating provides improved cover seals.

All reference frequencies for the receiver as well as all other station frequencies and time bases are derived by synthesis techniques from a master rubidium oscillator with a one minute stability of 5 parts in 10^{12} and a one year stability of 5 parts in 10^{11}.

Master station control is supplied by the special purpose computer shown in Fig. 10. This digital device currently controls transmit-receive cycles, range gates, and ranging codes, and is expected to take over such functions as automatic checkout and monitoring.

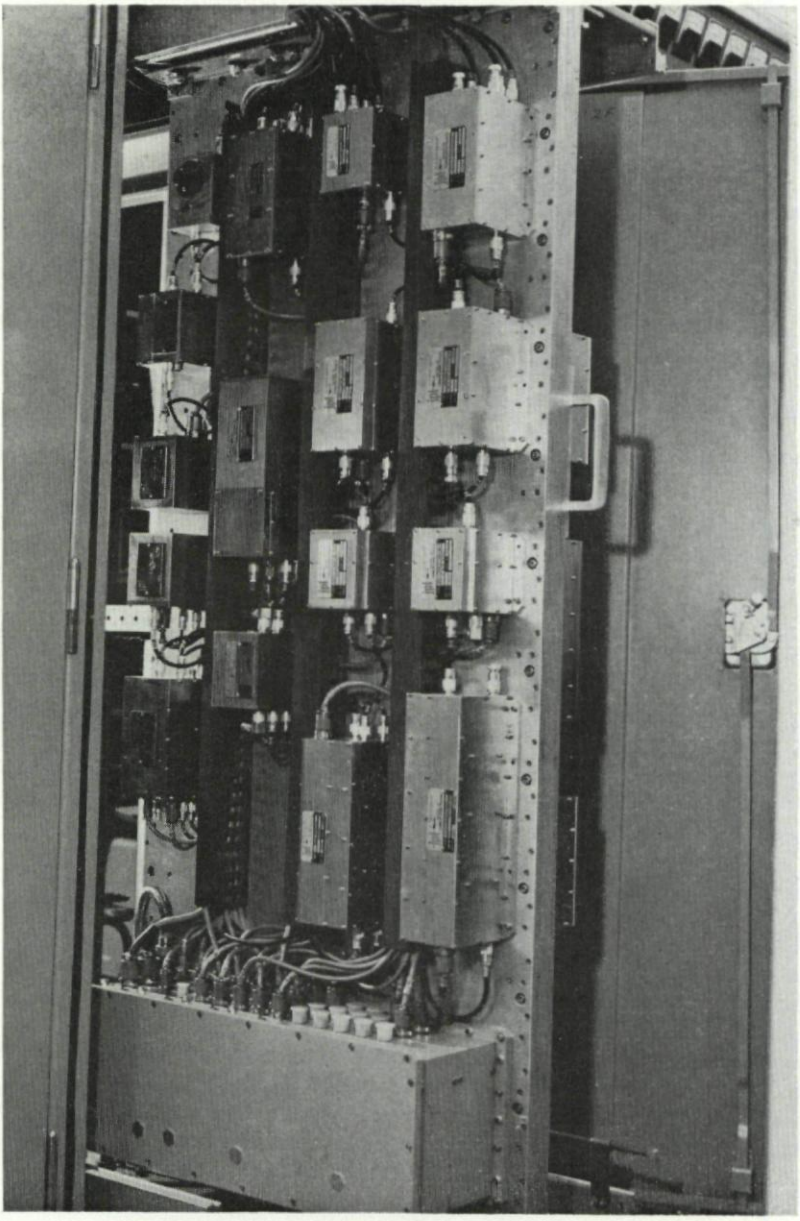


FIG. 9. *Receiver Mod IV interior.*

The programmed local oscillator (see Fig. 11), provides a local oscillator injection signal derived from an ephemeris tape and keeps the receiver tuned to the correct Doppler frequency during open loop operation. It is not necessary during locked loop tracking operations.

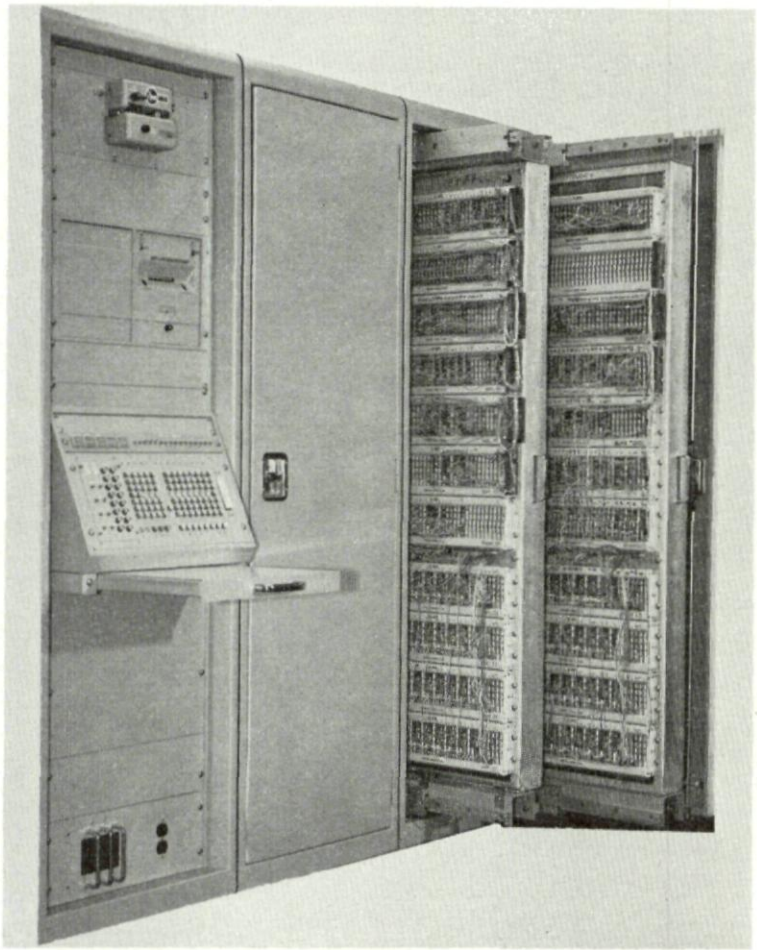


FIG. 10. *Master controller Mod III.*

The receiver output signal is integrated and processed, and the power spectrum is plotted with the equipment shown in Fig. 12. The transmitter exciter and power amplifier, which will be described in detail later complete the system.

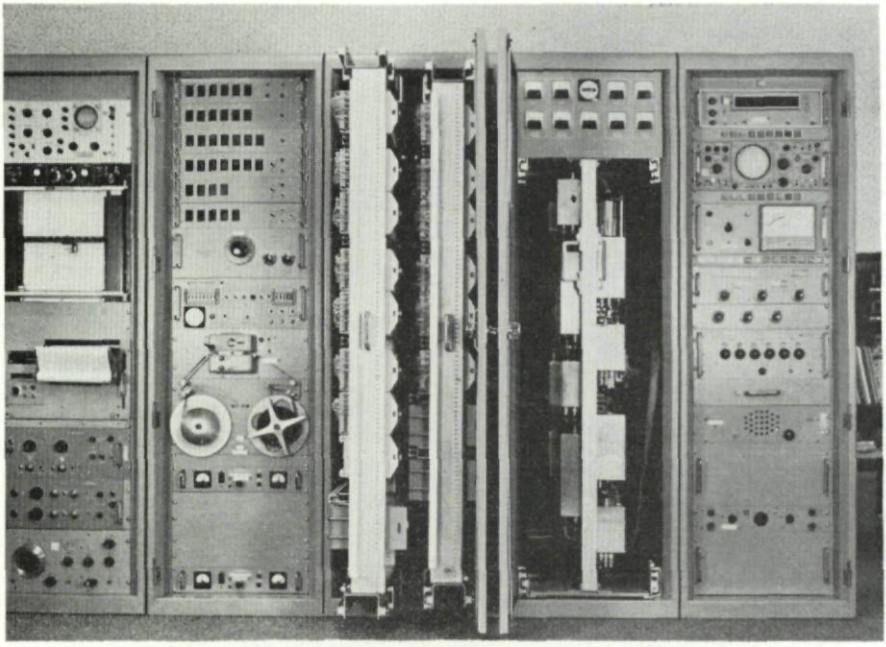


FIG. 11. *Programmed local oscillator.*

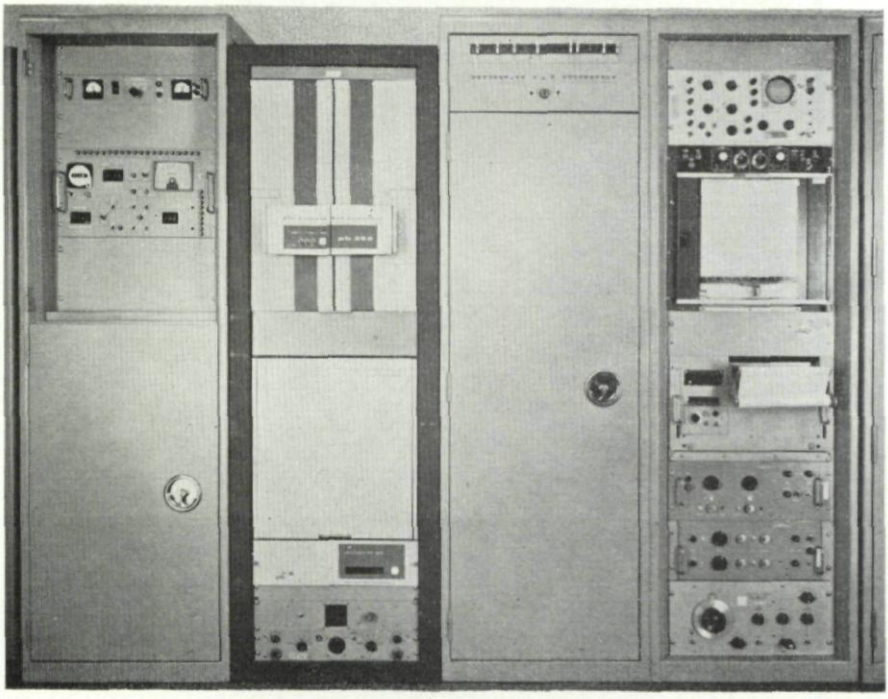


FIG. 12. *Data processing equipment.*

4. SYSTEM CAPABILITY

Since the system described is an advanced development system, there is a constant interest in improving the capabilities and range. Figure 13 shows graphically the increased capability of the JPL planetary radar since it was first operational in mid 1960 using a 10-kW transmitter, a receiver with a noise temperature of 1570°K and an 85-ft antenna. The installation of a ruby maser in early 1961 reduced the system noise temperature to 64°K. The transmitter power was raised to 13 kW, and improvements were made

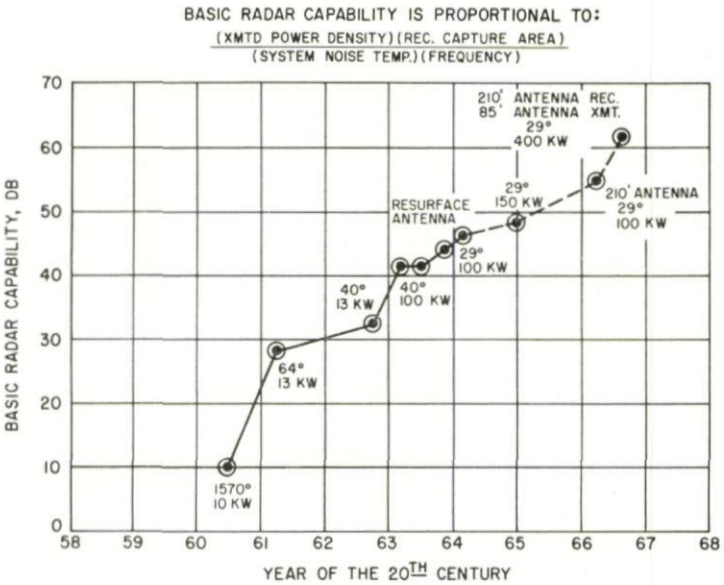


FIG. 13. *Radar system improvement chart.*

in the transmitting and receiving antenna feeds effecting a total improvement of 19.1 db. In late 1962, the change to a Cassegrain antenna feed system and the installation of a double-cavity maser improved the system capability by another 2.9 db. In early 1963, an improved two-cavity maser reduced the system noise temperature to 37°K providing a 0.3 db improvement. The 100 kW transmitter (used for the first time) increased the capability by 8.9 db. The 1.9 db increase in 1964 resulted from an improved antenna surface and the installation of a traveling-wave maser with a system noise temperature of 29°K. The projected improvement in 1965 and 1966 will result from use of the 210-ft antenna now under construction at Goldstone, California (8 db on transmit and 8 db on receive) and an increase in transmitter power to 400 kW (4.2 db).

The possibilities of increasing system capability are seen to consist of larger antennas, lower noise receivers, higher powered transmitters and

perhaps improved detection and signal processing techniques. The latter have been investigated intensively and present systems are near optimum for the state of the art. The 210-ft steerable antenna mentioned previously represents approximately the practical limit of construction. A maser with a noise temperature of 29°K does not leave room for much improvement in this area. Increased transmitter power remains as an attractive possibility, however.

High power alone is not enough in a CW radar transmitter. If this were the case, it might be more easily obtained by using a crossed field oscillator such as a magnetron instead of the exciter-linear accelerator power amplifier which is used. The requirements of such a transmitter are the following:

(1) High power
(2) Low incidental phase modulation
(3) Phase stability
(4) Amplitude stability
(5) Frequency stability
(6) Bandwidth
(7) Low harmonics
(8) Low noise
(9) Phase modulation
(10) Keying
(11) Linear capability

Low incidental phase modulation (or jitter) and medium term phase stability requirements are imposed by the phase sensitive system with which the transmitter is used. Incidental phase modulation may occur in the exciter oscillator or multipliers, or in the klystron amplifier. Changes in beam voltage on the klystron result in phase changes. In the type VA858 tubes used in this transmitter, the "pushing factor" is about 1° per 20 V and is a function of the electrical length of the klystron drift path. The objective on this system is a total phase jitter of less than 3° rms, which imposes a severe requirement on power supply filtering. Direct current is used on the tube heater to avoid modulation.

Amplitude stability of 0.1 db over one planetary transmit-receive cycle is desired. This again, in the klystron, is a function of power supply regulation. The exciter is designed for amplitude stability, but any instability is largely eliminated from the system by operating the klystron in a saturated condition.

Frequency stability is controlled entirely by the oscillator and as the basic signal is synthesized from the rubidium standard, it is held to 5 parts in 10^{11}, long term.

Bandwidth has never presented a problem in this type of operation. The exciter chain can be designed for any required bandwidth and the klystron exhibits a 3-db bandwidth of from 10 mc/s to 30 mc/s at 2388 mc/s (depending on the tuning). The maximum requirement to date has been 3 mc/s. The klystron could easily provide a bandwidth of 60 mc/s with a reduction in gain from 60 db to approximately 46 db.

Allowable harmonic radiation is limited by interference to other services. Originally, the transmitter was designed to include a harmonic filter. Subsequent measurements showed the second, third, and fourth harmonics to be down 42 db, 47 db and 72 db, respectively, which met all requirements without the use of the filter.

Noise in the klystron is approximately 10 mW in a 10-mc/s bandwidth with nominal beam voltage of 34 kV. This is 65 db below the transmitted power of 100 kW in a useful bandwidth of 3 mc/s and causes no problem on the transmitted signal. However this noise persists during receive cycles when the drive is removed from the klystron and poses a problem of isolation from the receiver. A low-noise transmitter is therefore desirable.

In many modes of operation, the ability to phase-modulate the transmitter with digital signals up to 1 mc/s and to $\pm 90^\circ$ is a requirement. On-off keying of the drive is also used and both of these functions are provided in the exciter chain to be covered later. Saturated operation is the usual mode, but specialized requirements for linear operation exist. For this reason some provision for linear operation is desirable. The Goldstone transmitter is capable of linear operation to 70-kW average power with not over 10% nonlinearity.

It can be seen from the above restrictions that the problem of providing high-power output with acceptable characteristics is formidable. The Goldstone transmitter, which is the subject of this Report, has been operational for 18 months at 100 kW continuous wave output and has been operated as high as 150 kW. In addition, it meets all the parameters listed in Fig. 12 and operates 12 to 20 hours per day, 7 days a week, with less than 3% unscheduled downtime. This represents a considerable improvement from the first operation in early 1963 when the mean time between failures was 17 minutes. System, component, and circuit redesign were required.

5. THE TRANSMITTER

As shown in Fig. 14, the transmitter consists basically of a power supply which converts the line voltage of 12,600 V, 3 phase, 60 c/s to direct current at up to 55 kV and 30 amp with a power limitation of 1 MW, for the klystron amplifier beam. The frequency synthesizer and the exciter provide an input signal to the 100-kW amplifier which provides approximately 60 db power gain. The transmitter control furnishes monitoring and control of all functions while 38 protective devices prevent damage to equipment by removing voltage in event of a malfunction. The liquid-to-air 1.5 MW heat exchanger is used to cool the amplifier, the power supply, and various auxiliaries to the transmitter.

The signal starts in the exciter which is shown in block diagram form in Fig. 15. The output of the rubidium oscillator, which is the basic station frequency standard, is synthesized to provide 31.84 mc/s. The coaxial relay is used to assure adequate isolation of the driving signal during receive periods. This is followed by the solid state keyer which can handle on-off keying from dc to 1 mc/s. The X5 varactor multiplier is followed by a phase

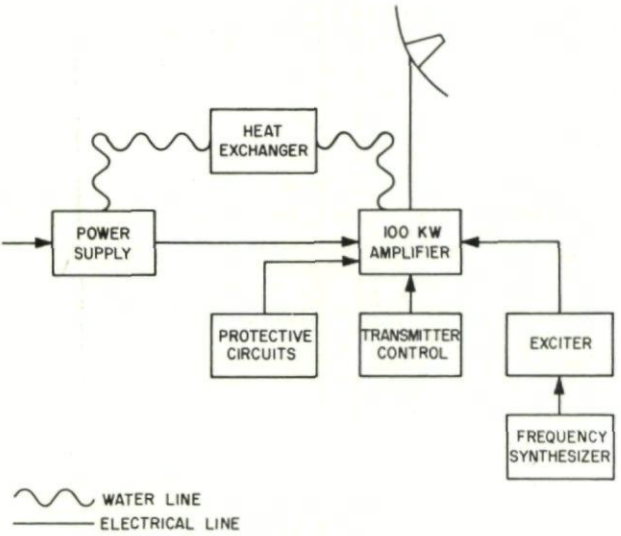


FIG. 14. *Transmitter block diagram.*

modulator. The only type of modulating signal encountered in normal modes of operation is digital square waves, so a special case is presented to this modulator. The ability to handle complex waves is not required but a high order of stability is necessary. The modulating signal operates two VHF solid state switches to vectorially sum a reference and a phase shifted signal to produce a phase modulation up to $\pm 90°$ at the transmitter output. The modulating signal is therefore relieved of any rigid amplitude stability requirements. Using symmetrical modulating waveforms, this system produces up to 60 db of carrier suppression at $\pm 90°$ shift.

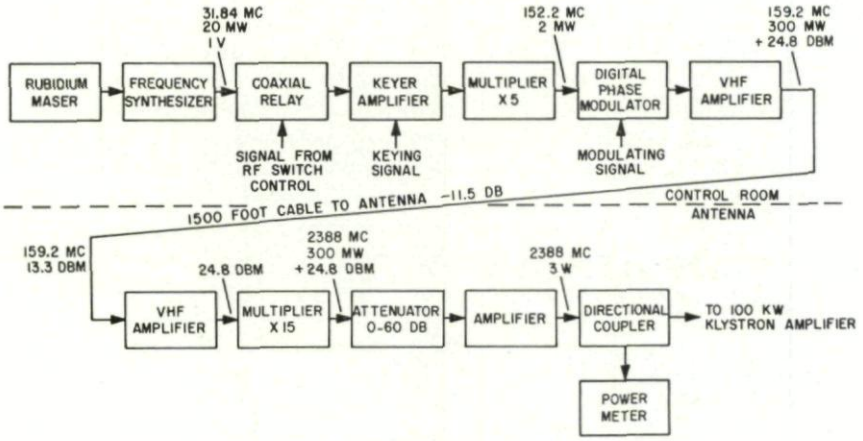


FIG. 15. *Exciter block diagram.*

The signal is amplified to 300 mW and sent through a 1500 ft cable to the transmitter compartment on the 85-ft antenna. The loss in this cable is 11.5 db, over half of which occurs in the antenna wrapup and terminal cables where the use of low-loss foam-type cable is not practical. A second VHF amplifier in the antenna compensates for the cable losses to drive the X15 solid state multiplier. The 0 to 60 db attenuator and the triode cavity amplifier provide a drive signal to the klystron from essentially 0 to 3 W. The directional coupler and power meter indicate the drive level on the antenna and in two remote locations.

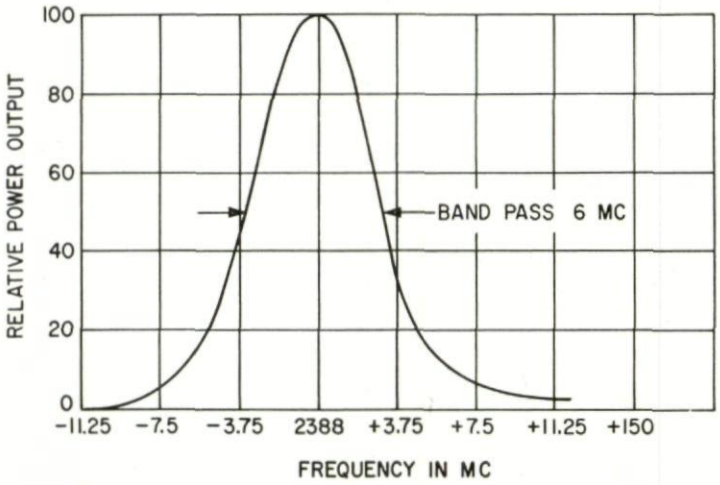


FIG. 16. *Exciter frequency response.*

Since the klystron gain is 60 db in the usual tuning condition, only 100 mW of drive is required. It is possible to obtain up to 2 W from a varactor multiplier at this frequency and plans call for replacing the present X15 multiplier with a unit capable of this output. The triode amplifier will therefore be unnecessary, and a major maintenance problem will be eliminated.

Figure 16 shows the bandpass of the exciter to be 6 Mc/s at the 3-db points. This exceeds all system requirements by a factor of 2. The control room portion of the exciter is shown in Fig. 17. The left cabinet is the 100-kW amplifier remote control with the rubidium standard and instrumentation in the center. The sliding plate at the right mounts the exciter proper. Figure 18 shows the antenna mounted exciter modules with the box cover removed. The elaborate double shielding is required to avoid danger of oscillation in the high-gain klystron. It also provides weatherproofing. Note that all cooling is by conduction to the box sides.

5.1 *The Power Amplifier*

The klystron power amplifier is the heart of the highpowered CW transmitter and is the item which required the most development work. Figure 19

shows the amplifier (with the top weather cover removed) mounted in the cage behind the antenna. Since the cage moves with the antenna in azimuth and elevation, the amplifier must operate in any attitude. The required 50 gal/min of cooling water is handled in azimuth movement by a rotary joint

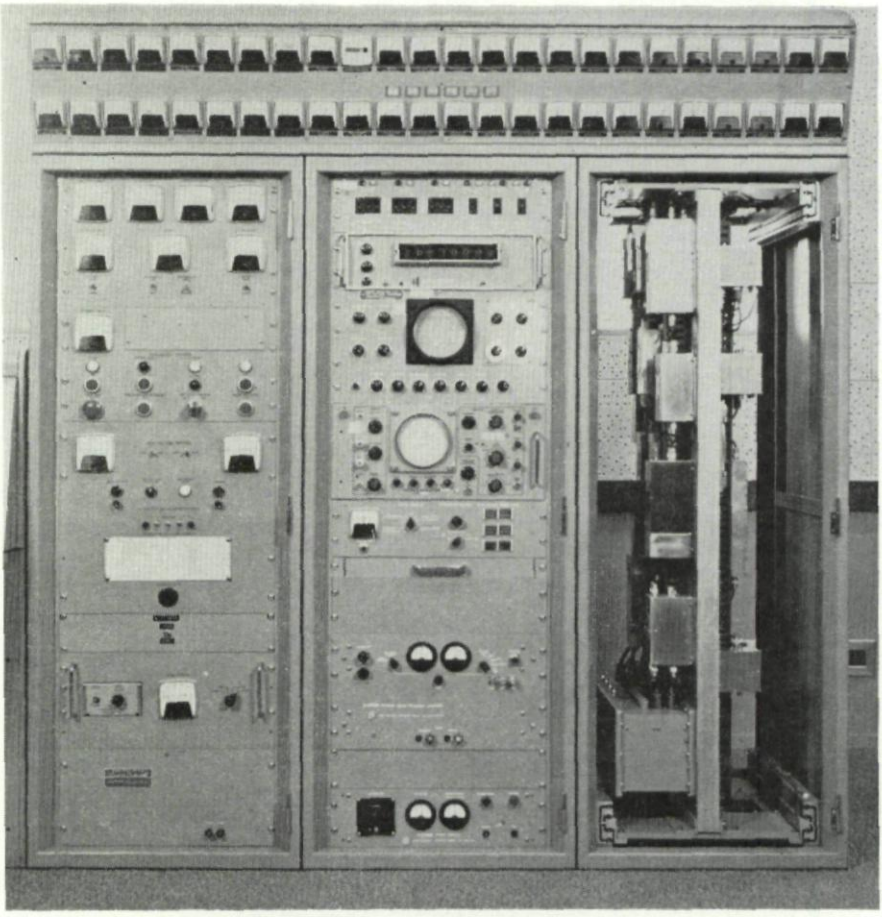


FIG. 17. *Exciter (control room).*

and in elevation by flexible hoses. A total of 4 cables containing 4 coaxials and 46 pairs is conducted from the ground to the transmitter via a wrapup for azimuth movement and flexible cables for elevation. Figure 20 is a block diagram of the power amplifier. The signal from the exciter passes through two series-connected crystal diode switches and through a ferrite isolator to reduce mismatch to the first cavity of the power klystron. This part of the circuit is in 50-Ω coaxial cable. The function of the crystal switches is to remove the drive in the event of a waveguide arc as indicated by a signal

from the arc detector, or a rapid increase in back power. These switches provide 60 db of isolation with an operating time of 5 μs for protection of the klystron. The 100-kW output is in WR430 waveguide to the waveguide switch, which selects either the antenna or the RF water load. This load will handle the full output continuously and is used for testing and adjustment of the transmitter as well as the basic standard for calibrating the power meters.

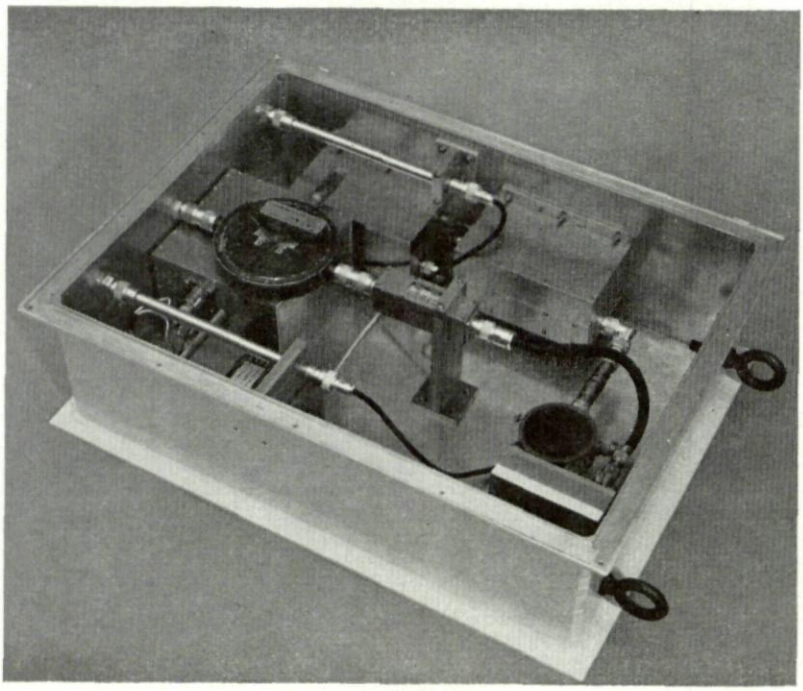


FIG. 18. *Exciter (antenna).*

The input and output water temperatures and the flow rate are monitored and power is computed by calorimetric methods.

The beam power supply input is normally 32 kV at 8.5 amp with the positive grounded and the negative high voltage to the cathode. At 100 kW out, this indicates an efficiency of 37%, which is affected a few percent by the tuning of the cavities. The filament supply of approximately 10 V at 10 amp is connected to the cathode and therefore operates at 32 kV above ground. The focusing magnet coils are water-cooled and are in series across a single power supply which, at best focusing adjustment, delivers approximately 500 V at 9 amp. The Vacion gauge gives a continuous indication of the vacuum in the tube and is particularly useful in indicating return to a safe operating condition after an outgassing in the tube. In addition, the Vacion assists in pumping down the tube. Forward and back power couplers

indicate power on local and remote meters, and in addition, the back power coupler provides tube protection as previously mentioned.

The four cavities of this klystron are tuned by reversible gear motors and tuning may be accomplished either locally or remotely. After a tube is placed in service, there is little occasion to change the tuning and it is thought that the motor tuning could be replaced with manual controls on future

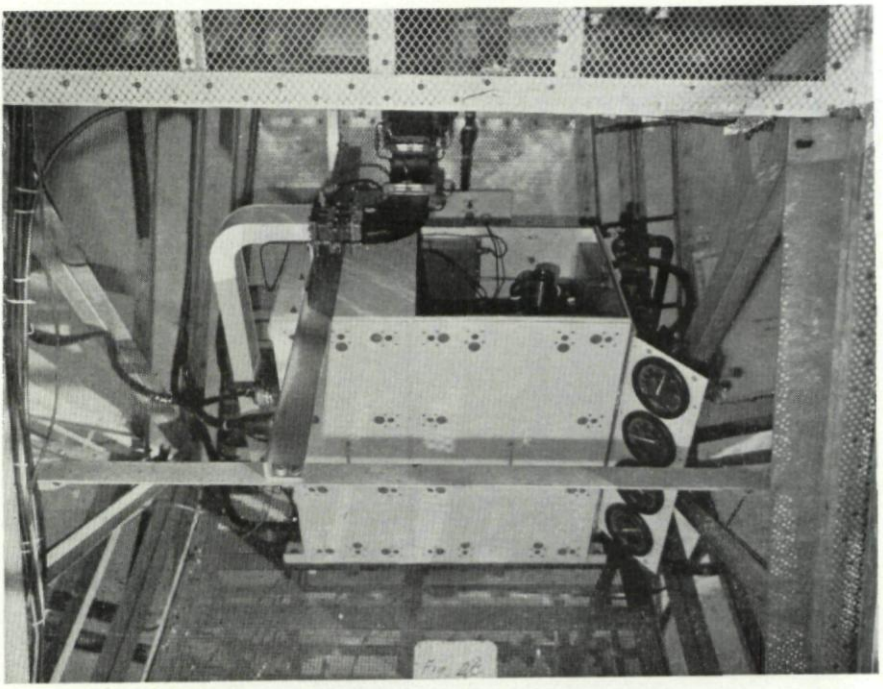


FIG. 19. 100-kW *amplifier in antenna.*

tubes. Further, future tube designs may well be fixed-tuned with considerable simplification and no compromise in performance.

Figure 21 is a view of the power amplifier on the ground with the weathertight covers removed. The gauges to the right indicate water flow, pressure, and temperature. The klystron and its magnet are in the right compartment with the two water hoses attached to the inlet and outlet fittings. The waveguide output is at the left. Because of the high powers involved, only oxygen-free copper guide is used to hold losses to a minimum and even this guide requires water cooling.

The main control station for the transmitter, which provides switching functions, instrumentation, remote tuning, focusing control, and fault indication, is shown in Fig. 22. This is located in the transmitter building near the antenna, and control may be maintained from here or from the remote panel in the main control building 1500 ft away.

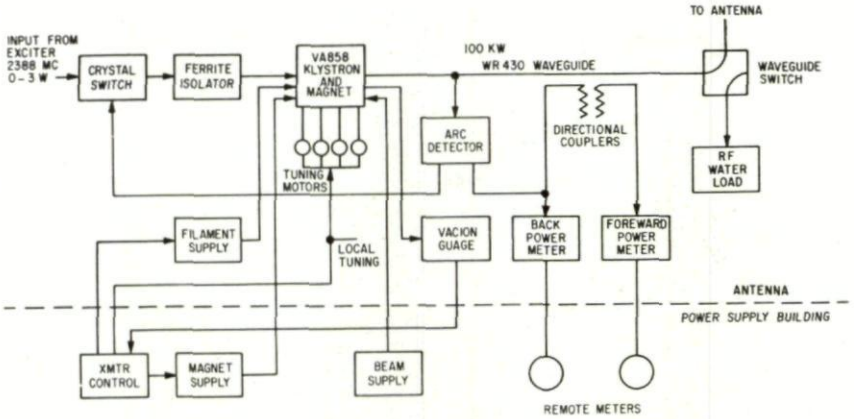


FIG. 20. 100-kW *amplifier block diagram.*

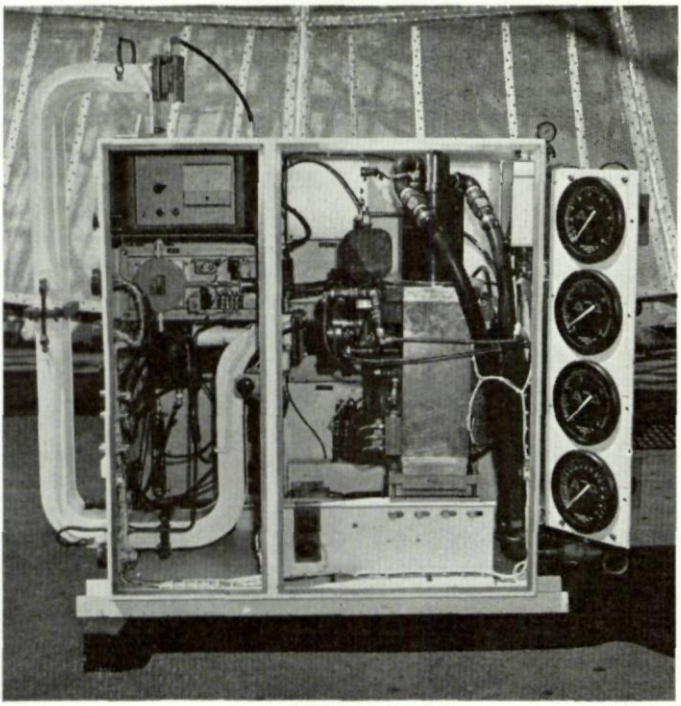


FIG. 21. 100-kW *amplifier (open).*

5.2 *The Klystron Tube*

The type VA858 klystron was developed specially for use in this transmitter, and is believed to be the highest-powered CW tube at this frequency in existence. It is 36 in. long by $10\frac{3}{4}$ in. diameter and weighs 135 lb. The

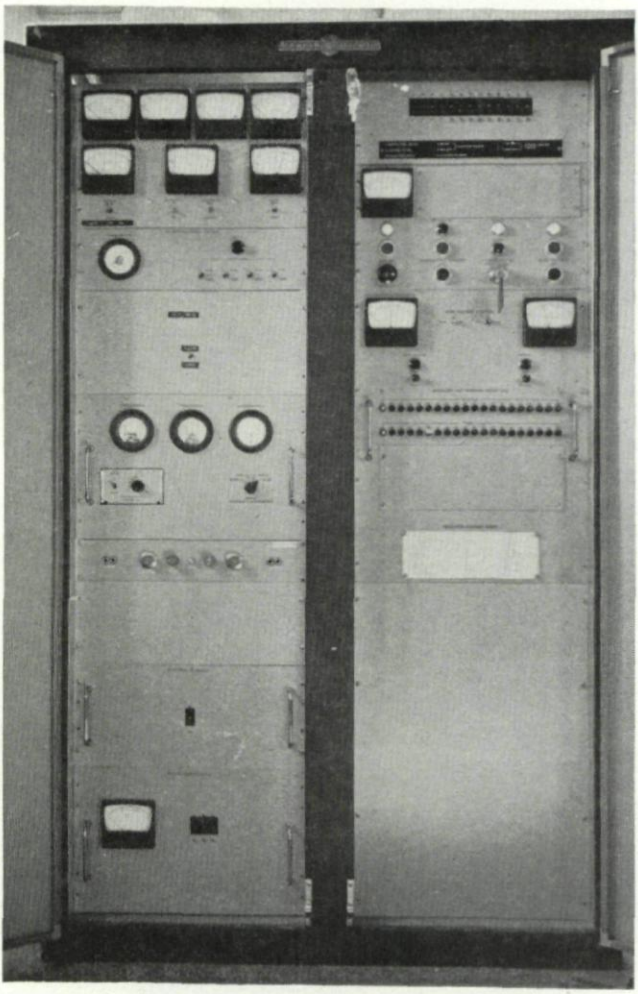


FIG. 22. *Transmitter control console.*

magnet weight is 425 lb. It is a four-cavity klystron and provides 46 to 60 db gain depending on the broadness of the tuning. Figure 23 is a view of the tube. Two klystrons were built on the original development and these tubes have been used for 18 months on planetary work. Three additional tubes have since been procured, but as yet have not been placed in service. The klystrons are satisfactory considering that they are the first tubes at this CW power level, and as such are highly developmental. However, they remain the limiting factor on power output and reliability of the transmitter, and do not average over 900 hr of service. A total of seven failures have occurred between the two tubes and after each failure the tubes have been rebuilt. Rebuilding costs average about half the cost of a new tube. One failure

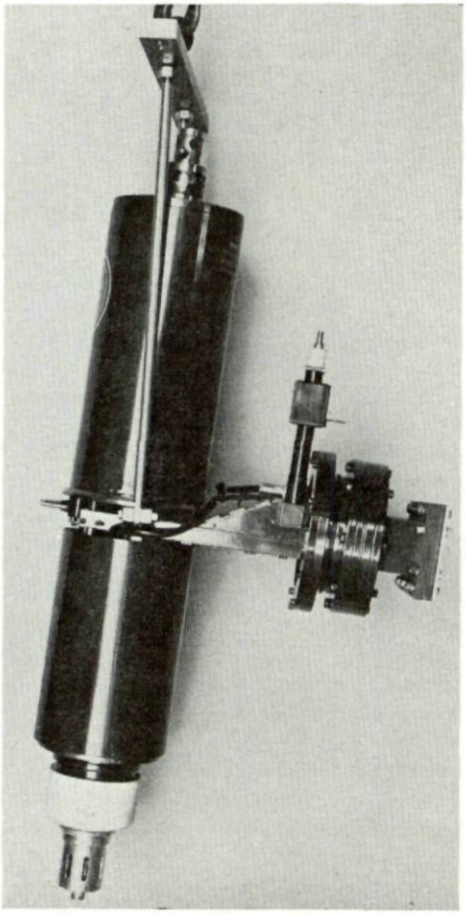


FIG. 23. 100-kW *klystron tube.*

occurred as a result of a few plugged water passages in the collector. A filter with smaller openings than the collector has been installed to avoid this problem. One output window cracked from heat during a high-power test at 178 kW which is, of course, 78 kW above rating. A nitrogen-forced cooling system is under consideration to remove this limitation. A third failure of this tube was caused by the manufacturer failing to properly clean foreign material from the tube before sealing. The second tube suffered three failures resulting from malfunctions of external equipments. This tube has two additional failures resulting from a cracked collector and a leak in a ceramic-metal seal.

The problem of collector erosion from the high-intensity beam is a major limitation on life. The $\frac{3}{8}$ in. thick copper collector walls can be eroded away to a point where water leaks occur. Future designs will use a collector with

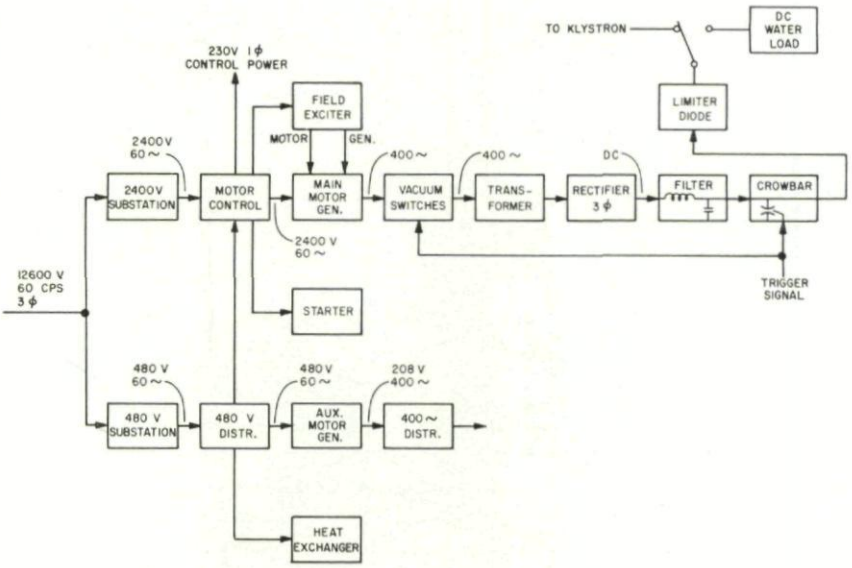


FIG. 24. *Power supply block diagram.*

a silver coated interior and shielded magnet coils to provide better beam dispersion in an effort to reduce this problem. Most damage results from high beam energy per unit area due to poor defocusing.

5.3 *Beam Power Supply*

The third major component of the transmitter is the beam power-supply. This supply will deliver up to 55,000 V and up to 30 amp of direct current

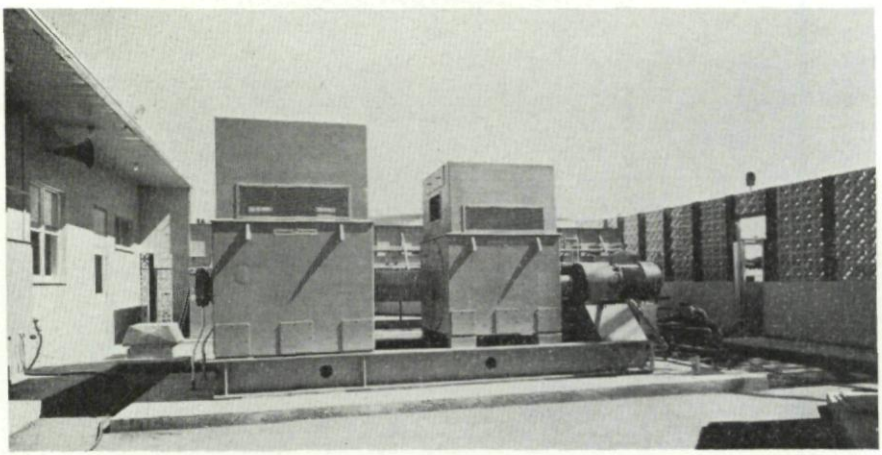


FIG. 25. *Main motor generator.*

with a power limitation of 1 MW. It is obviously overpowered for the 350 kW input requirement of the 100 kW transmitter and was designed with a view to future power increase. The supply is shown in block diagram form in Fig. 24. Power at 12,600 V, 3 phase, 60 c/s is supplied to two separate substations from a commercial line which is underground for the last mile because of the presence of strong RF fields. The 2400 V substation supplies the main motor generator only, while all auxiliaries are supplied from the 480 V substation. A 75 kW, 400 c/s, 3-phase motor generator operated from this line supplies all auxiliary 400 c/s requirements. The output of the main motor generator at 400 c/s is stepped up in voltage in the transformer, rectified, and delivered to the load through a filter, crowbar and series-limiter diode at voltages adjustable from 3000 V to 55,000 V.

Referring to Fig. 25, the motor generator consists of a 1750 hp, 60 c/s, 1800 rpm synchronous motor directly connected to a 1100 kVA, 400 c/s, 3-phase alternator. At the right is the 150 hp induction type cranking motor used to bring the unit up to speed through a magnetic clutch. Output voltage of the DC supply is controlled manually or automatically by adjustment of the field of this alternator (which is supplied from a rotary exciter). This exciter also supplies the field of the synchronous motor and this is used to maintain an essentially unity power factor for the total load by over

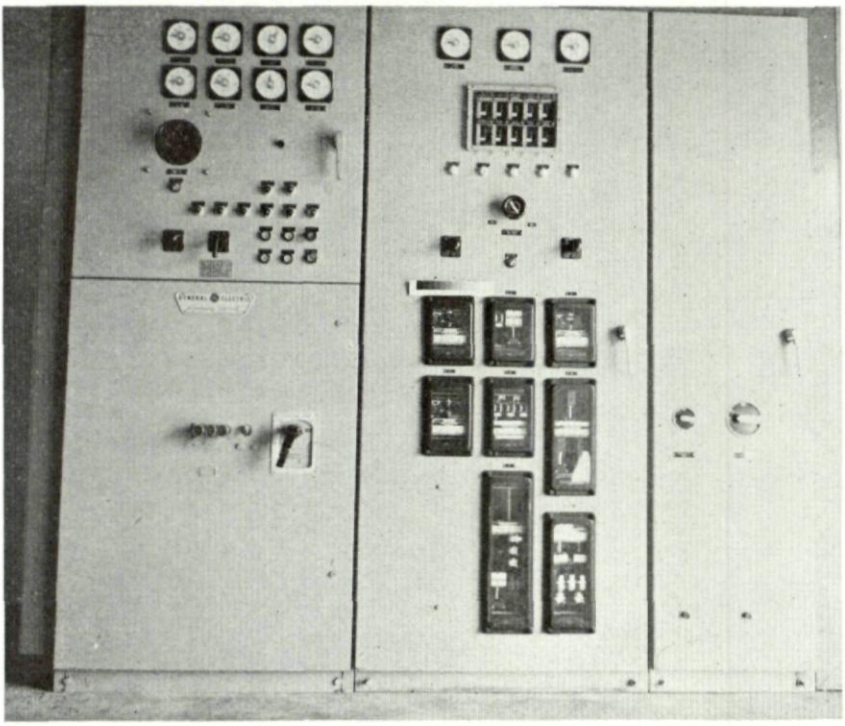


FIG. 26. *Motor generator control console.*

exciting the motor field. Control of the motor generator is accomplished in the cabinet shown in Fig. 26, which provides switching functions, over and under voltage and current protection, automatic and manual voltage regulation, synchronization, power factor control, and monitoring.

The use of a frequency changer (such as the motor generator) might seem unnecessary but actually provides technical and economic advantages which make it well worth while. It will be incorporated into future designs. It isolates the power line from a crowbar of the DC supply and greatly simplifies the line protection problem. It also isolates the supply from short duration line voltage fluctuations and transients due to the large inertia of its rotating components. The change from 60 to 400 c/s reduces all transformer sizes and costs, and by changing the ripple frequency from 360 to 2400 c/s, reduces filter size and cost. No failures other than normal start-up

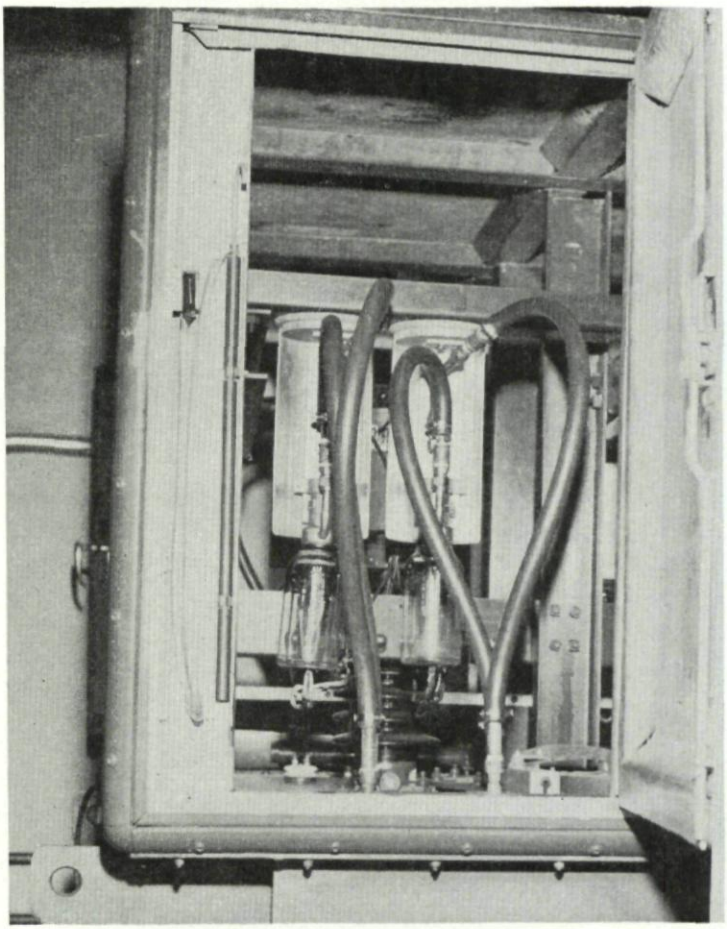


FIG. 27. *Transformer and rectifier.*

problems have occurred in this equipment and it is considered satisfactory.

The vacuum switches following the alternator are used to remove the residual field output of the alternator when it is desired to reduce the DC output to zero (as in case of a crowbar firing resulting from a fault). The vacuum relays were used instead of air or oil switches because of their fast opening time of 50 ms. The transformer-rectifier assembly is located in a concrete vault for safety. Figure 27 is a view of one corner of the transformer

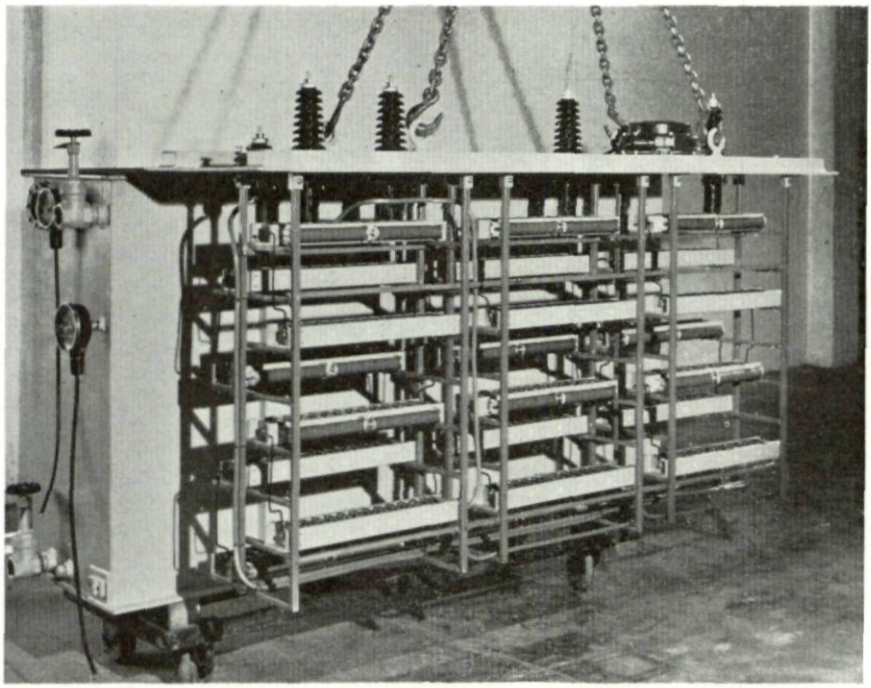


FIG. 28. *Silicon rectifier.*

and two of the twelve rectifier tubes used in the 3-phase, full wave bridge. The small size of the vault precludes a more inclusive photograph. The lead-lined X-ray safety door to the rectifier is open. The water hoses cooling the tubes are in the foreground and the object at the left of the door is the safety grounding rod. The transformer contains the high-voltage winding and filament windings for the rectifier tubes. It is oil-immersed and water-cooled. These tubes were a major source of trouble during the early months of operation and their life in some cases was only minutes. Under full voltage operation, the peak inverse voltage across the tubes is 57,750 V. Even at 32 kV DC out, the piv is 33,600 V. This was within the published ratings of the tubes, but both internal arcs, which destroyed the tubes, and external arcs, which crowbarred the system off the air, were common. An intensive program of redesign was undertaken and the present tubes have

operated one year without arcs or failures. However, the use of silicon rectifier stacks is considered superior to thermionic diodes from the standpoint of reliability, reduced voltage drop, increased efficiency, reduced cooling requirements, and elimination of filament supplies. An oil-immersed, water-cooled silicon stack has therefore been manufactured and is shown in Fig. 28 removed from its tank. This unit will be installed during the next availability of the system and future designs will use this type of rectifier.

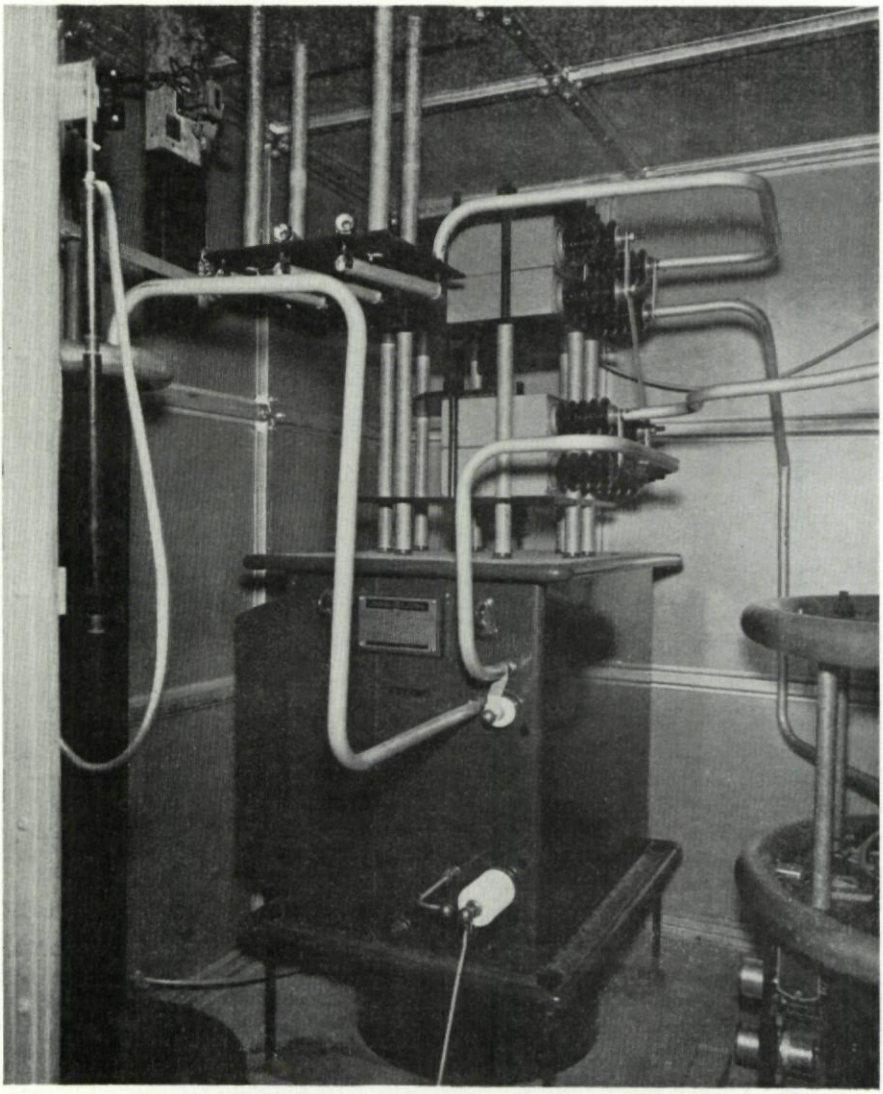


FIG. 29. *Filter.*

The output of the rectifier is conducted from the vault to an enclosure containing the filter, crowbar, crowbar logic circuits, and an automatic shorting bar. The filter, consisting of a 1-h choke and a 0.42 μF output capacitor, is shown in Fig. 29. The output ripple under full load is less than 0.05%.

An interesting phenomenon occurred when the system was first placed in service. The high-voltage cable is a double-shielded coaxial type with

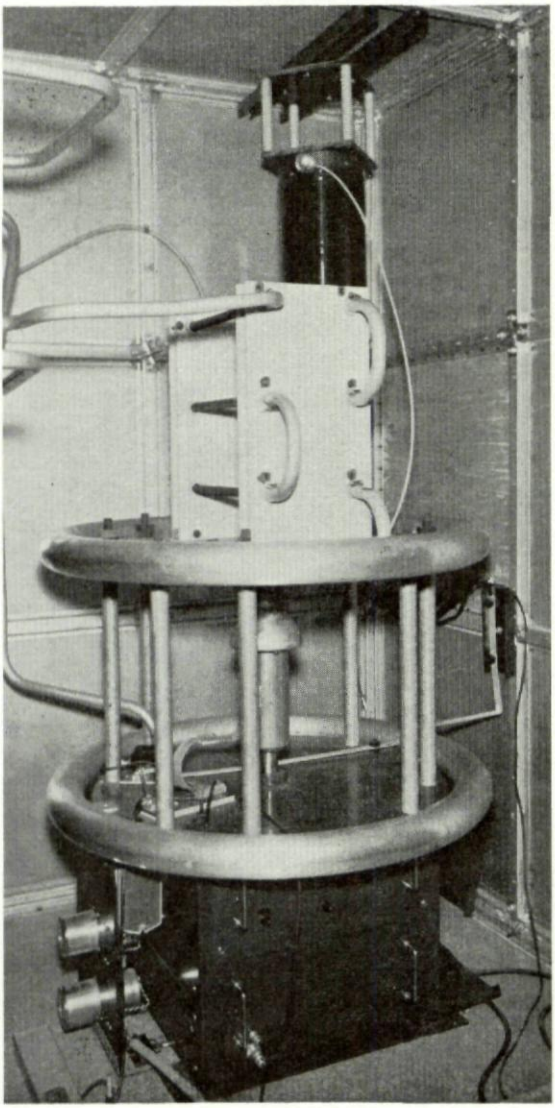


FIG. 30. *Crowbar.*

47 $\mu\mu$F/ft capacity. The capacity of the cable from the rectifier to the choke resonated the circuit at 2400 c/s (resulting in very high ripple). The addition of a 0.06 μF input capacitor shifted the resonance and solved the problem.

Also in the compartment with the filter is the crowbar protective device. This consists of a hemisphere gap (shown in Fig. 30). The high voltage out of the filter appears across this gap with the lower hemisphere at ground potential. The gap opening is controlled by a servo which increases the gap with increasing DC voltage. Located in the lower half is a spark plug which initiates a discharge across the gap when high klystron body current, a step function of beam current, or a rectifier tube arc is sensed. This discharges the stored energy of the filtter across the crowbar, rather than across the fault, preventing equipment damage. The operating time of the crowbar is 5 μs. At the time the gap is fired, the contactors to the alternator field are

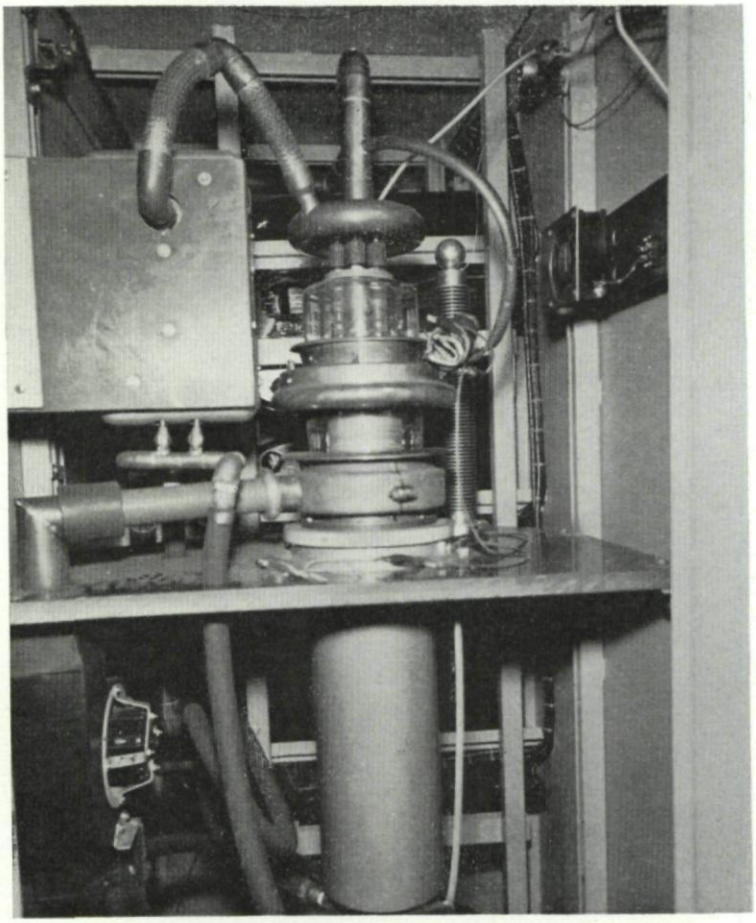


FIG. 31. *Series limiter diode.*

dropped and 40 ms later the vacuum relays between the alternator and transformer open. This removes the alternator residual field voltage and drops the system voltage to zero.

An additional klystron protective device is the series-limiter diode shown in Fig. 31. This is a 175 kW plate dissipation triode connected as a diode in series with the high-voltage lead from the crowbar to the klystron. The filament voltage is controlled to provide emission limiting of the pass current to 150% of normal, thus preventing high peak discharge currents in event of a klystron arc. While this device is effective, it has several drawbacks. It

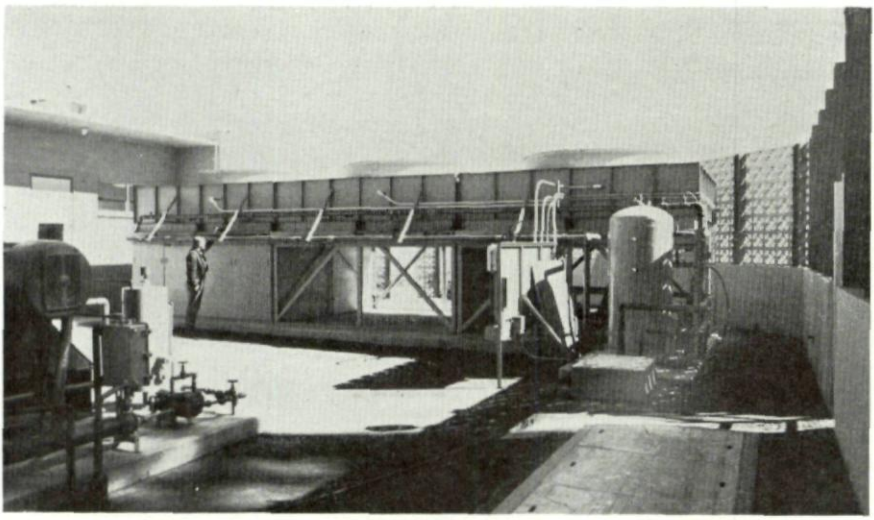


FIG. 32. *Heat exchanger.*

does not protect against the stored energy in the 17,600 $\mu\mu$F capacity of the cable to the antenna, which at 32 kV amounts to 8.7 j. The voltage drop across the diode at 9 amp is approximately 10,000 V, resulting in a 90-W power loss. In addition the filament of this tube consumes up to 21,600 W. If higher power operation is attempted, the supply does not have sufficient voltage output to afford this 10 kV loss. For these reasons, the diode limiter will be replaced by a 10-Ω surge resistor at the klystron. This will limit the energy in the tube arc to 2.3 j, which is 30% of the unlimited value and is adequate protection.

Because of high ambient temperature and frequent sand conditions on the desert, air cooling is generally not satisfactory. The only air cooling used is a small blower on the klystron cathode seal. The klystron, motor generator clutch, transformer, rectifier, magnet coil, series limiter and its filament supply are all cooled by distilled and deionized water from the main heat exchanger shown in Fig. 32. This unit has a capability of 1.5 MW in ambient air of -32°C to $+57$°C, with a water output temperature not to exceed

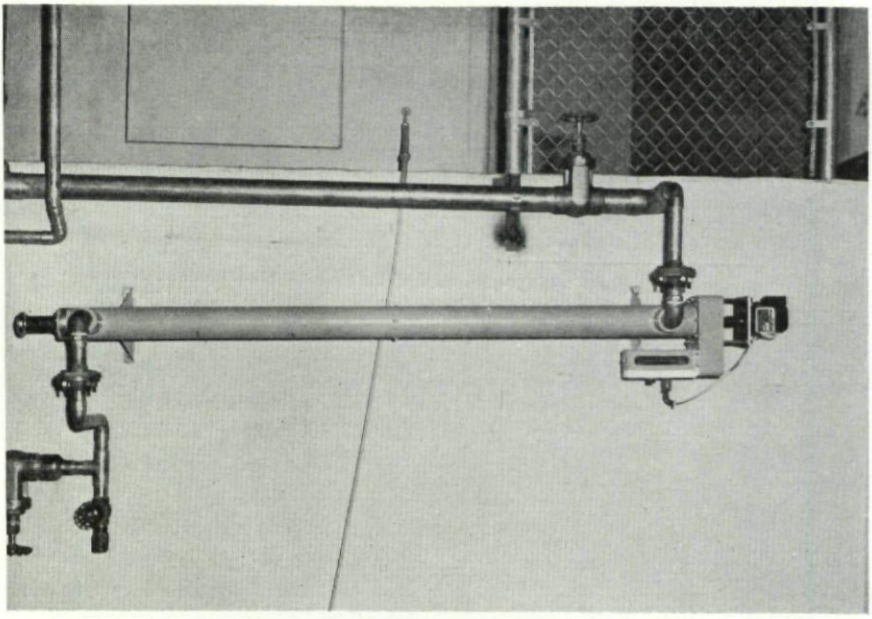


FIG. 33. *DC water load.*

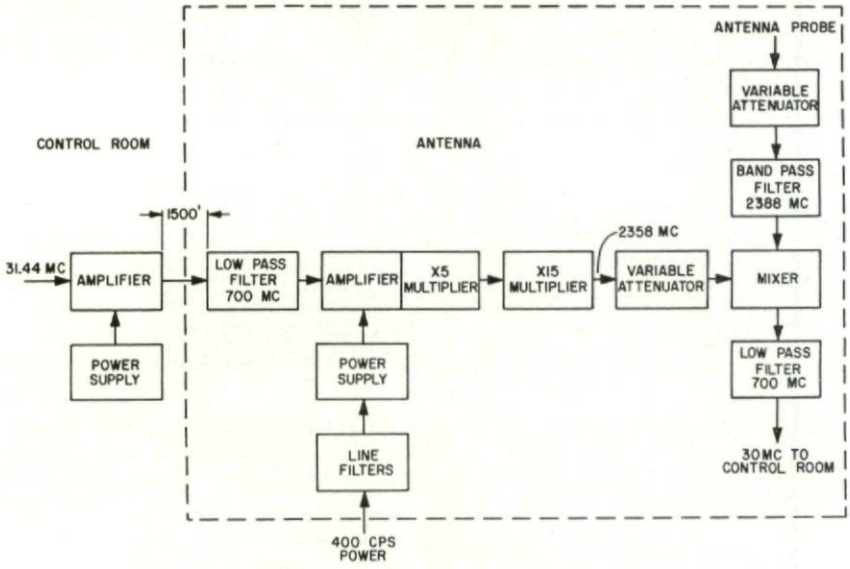


FIG. 34. *Spectrum analyzer block diagram.*

62°C. The pressure is 165 psi at the heat exchanger and 140 psi on the antenna at a flow rate up to 300 gal/min. There are three radiators cooled by three 10 ft fans. Temperature is automatically controlled by vane type air shutters. When a shutter is $\frac{1}{4}$ open, the associated fan shuts off and starts again when the shutter opens beyond the $\frac{1}{4}$ point. The four circulating pumps are manually controlled. A 30 kW heater is used to prevent freezing under extreme conditions and covers can be installed on the radiators for shut-down. In the picture, the covers are installed on the left radiator. A still and deionizer are being installed to prepare make-up water on the site. The heat exchanger is satisfactory and no problems exist with it.

Certain accessory equipment is used which is not required for operation of the system, but which is useful for testing and maintenance. A DC water load, Fig. 33, is used for testing the power supply. It is mounted high on a wall for safety and is remotely controlled. Heat is dissipated by using the cooling water as a conductor. The unit is satisfactory at the 350 kW level, but problems with insultator failure and internal arcing have been experienced at 1 MW. A new load is under design at JPL to handle the full power.

In order to monitor the transmitted spectrum in the control room, a spectrum analyzer of the swept frequency type is used. The input to this unit is centered on 30 mc/s. Figure 34 is a block diagram of the antenna-mounted weatherproof box which heterodynes 2388 mc/s to 30 mc/s for this purpose. The 31.44 mc/s synthesized from the rubidium standard is sent to the antenna and multiplied $\times 75$ in solid state multipliers to 2358 mc/s. This is mixed with the 2388 mc/s obtained from a short probe in the antenna, and the resulting 30 mc/s returned to the control room for display on the spectrum analyzer.

6. PROTECTION

There is extensive equipment protection provided throughout the system. Various devices have countless times prevented damage to the klystron or other components. Indeed, it has been suggested that it is statistically remarkable that the transmitter has operated at all with this multiplicity of cutoffs. In addition to the waveguide arc and back power sensors, there are 38 additional sensors which remove primary power from the transformer. These are low water flow, over current and under current, over voltage and under voltage, temperature, cooling air and high-voltage compartment interlocks. In the early months of operation, extreme problems were encountered with the actuation of the logic circuits of these devices by extraneous electrical signals. As an example, the operation of the klystron tuning motor switches on the magnet power supply Variac drive motor triggered one or more protective devices. There was even an indication that the pulse from the room light switch could cause a false operation. This has been eliminated by an extensive program of shielding and filtering leads and bypassing of spark-producing devices.

When super high-power was first considered, a major concern at JPL was for personnel safety. This continues to receive the daily attention of the engineers and operating crews. The value of the extensive precautions

observed has been proven by the lack of accidents at the site to date. There are three potential hazards: RF radiation, X-rays and high voltage.

The allowable standards for radiation have been established at 1 mW/cm^2 continuously, and 10 mW/cm^2 for short periods not to exceed 1 hour out of 24. These are the most conservative standards known and compare with 10 mW/cm^2 continuously, which is commonly used. To insure that these limits are not exceeded, there are 10 hand-carried radiation monitors with direct reading scales on the site. All areas have been surveyed and safe areas are documented. All personnel entering the radiation area are issued with one of these monitors. The transmitter and antenna area is surrounded by a fence with flashing orange lights when the high voltage is on, and flashing red lights when the transmitter drive is on. When any of these lights are on, only approved transmitter personnel are admitted inside the fence. This regulation is enforced by a guard on the gate. There is an automatic cutoff which shuts down the transmitter if the antenna is lowered below 7 degrees, which prevents the beam from striking buildings or personnel areas. A safety key switch on the antenna permits personnel to remove the key, thereby preventing the application of high voltage to the transmitter. In addition, there are key switches on the control panels in the transmitter building and the control room. All of these keys must be in place in order for the transmitter to operate. The only duplicate keys are kept at JPL under lock. Red "panic buttons" located in the klystron amplifier box, outside the box, and on both control panels will turn the transmitter off. In addition, an announcement is made over the public address system before the transmitter is energized.

The only unsafe areas are in the main beam of the antenna. Levels around the klystron and behind the parabolic reflector do not exceed 4 mW/cm^2, and most areas are below this. No operating area shows a field which will give an indication on the monitors with a threshold of 0.1 mW/cm^2

Soft X-rays are generated by the klystron and hard X-rays by the rectifier tubes. The latter will be eliminated with the installation of the silicon rectifier. The safe level of X-ray exposure has been established at 2.5 mr/hr, not to exceed 100 mr/wk. The level at the klystron cathode is 9 mr/hr and with the covers in place, 2 mr/hr. At a distance of 2 ft with the covers on, there is no measurable X-ray radiation. The rectifier tubes are enclosed in a lead-lined cabinet outside of which there is no measurable radiation. The cabinet doors are interlocked with the high voltage so no measurement has been made with the doors open. An X-ray meter is on site for survey or personnel use and all possible danger areas have been surveyed. Personnel working for extended periods in any possible danger area are required to wear film badges. In addition, the fence and key switch system described above also protects against X-rays.

The principal high-voltage danger is from the beam supply which can be as high as 55,000 V, but danger also exists in the 60 c/s and 400 c/s circuits. All compartments and cabinets containing high voltage have door interlocks which will turn the system off if any door is opened. Shorting chains or "dead man sticks" are located at all high-voltage points, and these are

required to be in place before any work is done on this equipment. All control panels are dead front with no exposed voltage points. Warning signs are prominently displayed. In addition to the "panic buttons" described above, there are cord-mounted buttons in the transformer-rectifier vault and in the crowbar compartment. The key switch system and personnel control fence are a major part of the high-voltage safety plan.

The most valuable safety device is a careful man, and the site personnel have received considerable safety indoctrination. JPL engineers are always on the alert to correct any unsafe condition or practice. While nothing can be completely foolproof, the transmitter is considered to represent a high order of safety engineering.

The experience gained with the 100 kW transmitter has placed the Jet Propulsion Laboratory in a position to proceed to even higher power. Developmental work is in progress on a klystron to utilize the full 1 MW capability of the power supply and deliver a minimum of 350 kW CW to the antenna. The problems will be many, but such a system will certainly be operable in the foreseeable future.

DISCUSSION

J. THRANE: I would like to ask the speaker if he could just give us one or two sentences about the arc detection device used and tell us how often it had to work.

W. S. BAUMGARTNER: How often it has to work is a function of how good is your system, in particular how good is your wave guide. When we first placed this on the air to hazard a guess, it worked about once every 60 seconds. We don't expect to have an operation out of the arc detector now more frequently than once in several months. The arc detector is a photo actuated device, the arc detector in this system actually has the photo cell mounted in a small aperture in a bend in the wave guide. The output of the photo cell is conducted through a transistor amplifier and control rectifier which initiates a discharge across the gap of the crowbar. We are using in another system a different arc detector in which rather than mounting the photo cell in the wave guide we are using some fibre objects which we can bend around at will like you might a soft rubber tube. One other thing that is incorporated in the wave guide which is highly essential is a small circuit to test the arc detector.

CHAPTER 24

TRANSMITTER RESEARCH IN THE U.K.

N. S. NICHOLLS
Royal Radar Establishment, Malvern, England

A survey is given of the in-house and extra-mural activities of the Electronic Group, R.R.E. in the field of devices and techniques for pulsed radar transmitters.

1. MICROWAVE TRANSMITTING VALVES

Transmitting valves envisaged for future pulsed radars include 'O' type beam tubes, such as klystrons, T.W.T.'s and hybrids, and magnetrons for use at the millimetre wavelengths. Crossed field amplifiers are considered to have promise in respect of compactness, lightness, efficiency, high perveance, and low A.M. to P.M. conversion, which may in due course make them the most attractive device for many purposes. However, there is not at present much activity outside the valve laboratories in this area.

Work on microwave transmitting valves will not be discussed at length in this paper because it is well covered by other conferences. Instead, the paper will concentrate mainly on research related to other aspects of radar transmitting equipment.

2. PULSE MODULATION

Beam-type microwave valves may be pulse modulated either by switching the H.T. supply or by means of a current controlling electrode built into the valve. The economy in switching of the latter arrangement (Fig. 1) is, however, often more than offset by the advantage of reduced voltages in the power supply circuits made possible, when the full H.T. supply is pulsed, by the use of a step-up pulse transformer. This is particularly significant when the peak power is above about a megawatt and the duty ratio is less than a few percent.

The use of a pulsed H.T. supply eases the demands on the electron gun of a microwave amplifier, and this is particularly valuable at high peak powers and at short wavelengths.

Provided that the desired sequence of pulses conforms approximately to a fairly simple pattern, the network-type pulse generator (Fig. 2) is usually the most economical modulator, even when a high degree of constancy of voltage on the transmitting valve is required during each pulse.

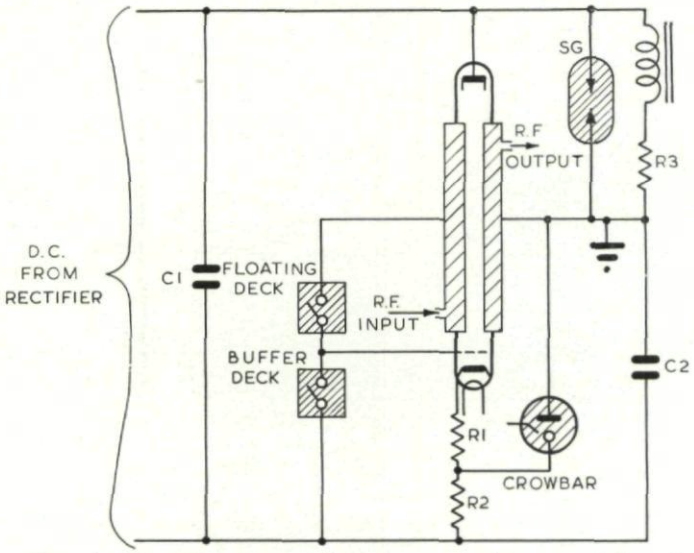


FIG. 1. *Circuit for pulse modulation by use of a beam control electrode*

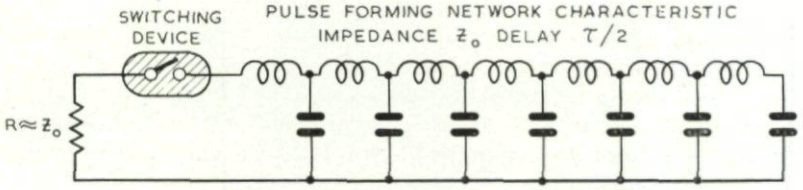


FIG. 2. *Network type pulse modulator*

3. HYDROGEN THYRATRONS

The most successful switching device for network-type modulators is the hydrogen thyratron. For small powers, envelopes of glass or ceramic are employed. The advantages of the latter are greater mechanical strength and better heat conduction. This type is well illustrated by the CX 1157 for operation at a maximum peak power of 3.5 MW and 3 kW mean at 20 kV. At higher powers, heat transfer from the grid assumes a major importance in design, and it appears advantageous to use a copper envelope. A higher voltage is also appropriate, and is obtained by the use of deuterium. The E.2829 is an example of this type, with peak and average power capabilities of 40 MW and 50 kW when operating at 40 kV. The envelope is $2\frac{1}{2}''$ diameter. The E.2816 is a 40 kV 3.5 A. rectifier version of this valve with a smaller cathode and more open grid.

The E.3057 is a larger device of the same basic design, with target ratings of 100 MW peak and 180 kW mean at 45 kV and is regarded as being about the most powerful thyratron feasible with this basic design.

For still higher powers, either a higher voltage may be employed, or passages for coolant must be incorporated within the envelope. A example of the former is the E.3058, with an average power capability of around 400 kW at 80 kV. Two voltage gaps in cascade are employed. The well established GHT6 is an example of the latter and has been operated at powers up to 200 MW peak and 350 kW average without distress.

With a potential life of around 20,000 hours, high precision of timing and overall efficiency from 98 to 99.5%, hydrogen thyratrons are regarded as capable of meeting substantially all forseeable requirements in network modulator switching at peak power levels above about one megawatt and seem unlikely to be superseded for a long time.

4. SOLID STATE MODULAR SWITCHES

The advantages of solid state devices—small size, ruggedness, long life and instant starting are offset in the pulse-modulation field by a comparatively

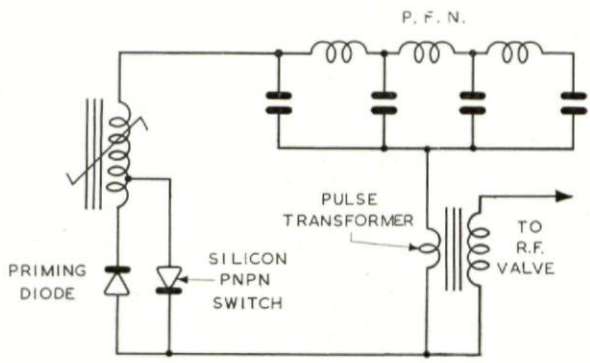


FIG. 3. *Use of semi-conductor switching device in conjunction with magnetic switch.*

low efficiency. At low powers, however, silicon thyristors have found favour. But to obtain the lowest power loss at the short pulse lengths commonly employed in low powered radars, not only must device design be optimised but it is also necessary to use an auxiliary magnetic switch and a means for priming. These may be conveniently combined in the manner shown in Fig. 3. Figure 4 shows the percentage power losses as a function of saturable reactor delay in a device operating at 400 V. 35 A. generating 0.3 μs pulses with a rise time of 0.06 μs.

5. AUXILIARY DEVICES FOR NETWORK-TYPE MODULATORS

The successful operation of network-type pulse modulators requires the use of various additional devices, such as inverse diodes, charging diodes, and power rectifiers.

For inverse diodes, where the duty is very light apart from occasional large surges when an arc occurs in the transmitting valve, semi-conductor diodes are attractive. Difficulty may be caused, however, by mal-distribution of inverse voltage due to minority carrier storage during the rapid oscillations

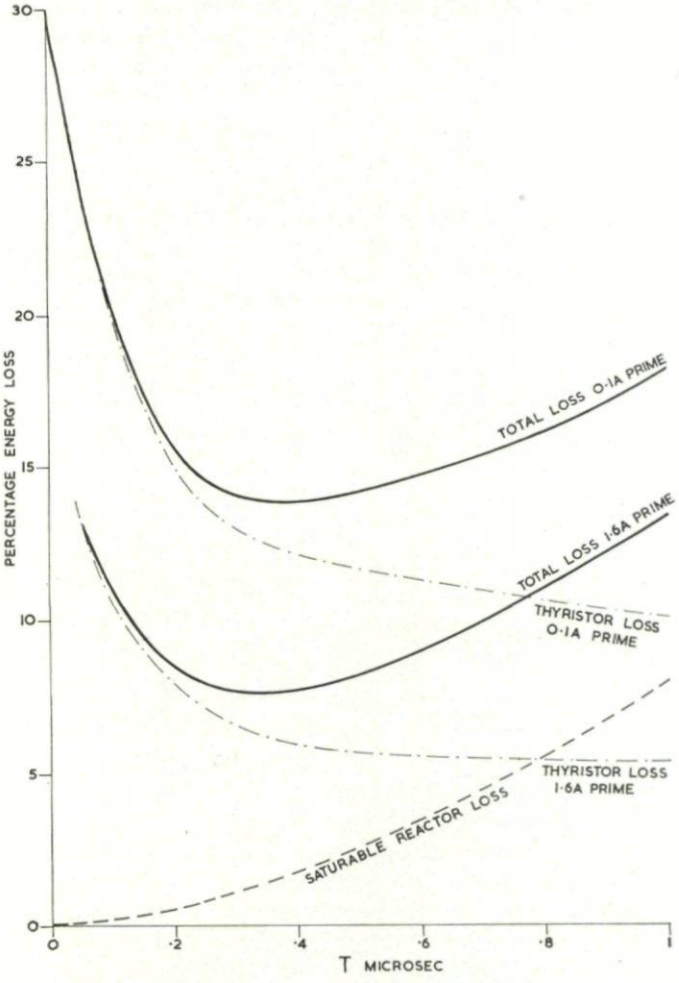


FIG. 4. *Percentage energy loss as function of reactor delay.*

which follow a normal pulse. A design to overcome this difficulty without the use of cumbersome sharing capacitors employs a mixed stack of high and low-voltage controlled avalanche silicon diodes. It is arranged by suitable choice of charge storage that sharp inverse transients are developed across the low voltage units which are better able to withstand the avalanche currents involved. However, since the post-pulse oscillations are small in amplitude,

the greater part of the voltage capability of the stack may be supplied by more economical high-voltage units.

H.T. supply for radar transmitters is best supplied either by semi-conductor or gas-filled controlled rectifiers. Controlled rectifiers provide voltage control and circuit breaking without the use of electro-mechanical devices which add weight and are deficient in life and reliability. At low powers,

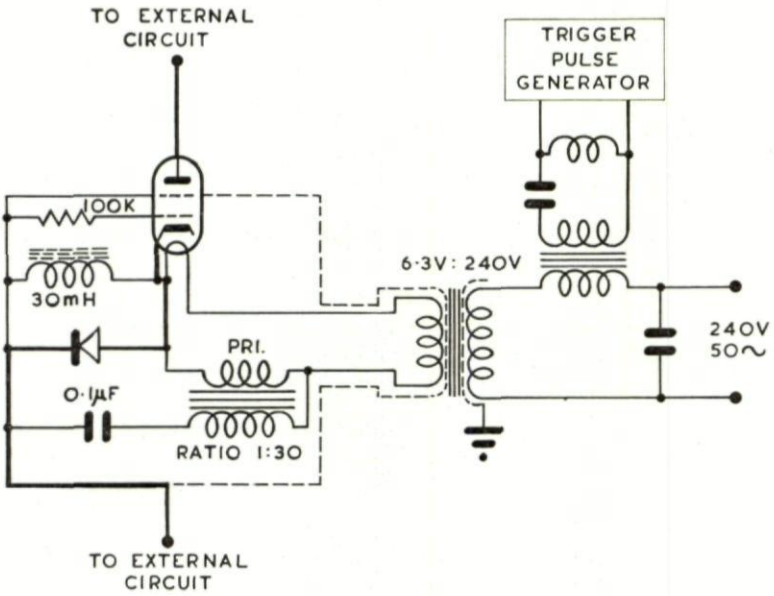


FIG. 5. *Thyratron circuit.*

control may be exercised by the use of back-to-back connected silicon controlled rectifiers in the H.T. transformer primary, but at power level above a few tens of kilowatts, deuterium thyratron rectifiers offer advantages in respect of power loss and compactness. For instance, an air convection cooled 30 kV 10 A. controlled rectifier unit using the E.2816 thyratron previously mentioned occupies a volume of 25 cu. ft. In this design, the rather large insulated heater transformers necessary for the thyratrons do double duty by also serving to convey the control pulses to the valves. The thyratrons are used in a "grounded–grid" circuit to screen their transistorized trigger circuits both from high-frequency interference and the effects of flashover. (Fig. 5.)

Greater compactness may be attained by imersion of the complete rectifier unit in an insulating and cooling liquid such as Fluorinated Ether. A volume of $1\frac{1}{2}$ cu. ft. suffices in this case for the 30 kV 10 A. unit.

AN EXPERIMENTAL 100 MW 400 mc/s TRANSMITTER

This installation was constructed in order to permit research on Travelling Wave Tubes with beam voltages up to 500 kV. The maximum

modulator power available is 250 MW in pulses of duration from 2 to 8 μs with a maximum duty ratio of 2×10^{-3}.

The modulator is in effect 6 separate modulators each using a GHT6 thyratron coupled to a common pulse-transformer giving a step up from 11 kV to 500 kV. The T.W.T. under test lies in a lead lined chamber beneath the modulator.

7. AN EXPERIMENTAL 20 MW 200 kW 400 mc/s TRANSMITTER

This installation is intended to form part of a radar for free electron scatter measurements. The transmitting valve is a klystron designed to operate with a beam voltage of up to 240 kV. This is supplied through a pulse-transformer from three modulators in parallel capable of a combined output of up to 40 MW peak, 450 kW mean with a pulse length of 200 μs maximum.

An unusual feature of this installation is the use of A.C. charging of the pulse forming networks, the three networks being charged sequentially from a 3 phase supply during each cycle of the supply. Output power is controlled by shifting the timing of the modulator relative to the supply waveform.

CHAPTER 25

PHASE SHIFTERS FOR PHASED ARRAY ANTENNAS*

H. A. HAIR
General Electric Company, Monmouth, U.S.A.

Fast switching microwave phase shifters are described for use in high power array radars. Drive control techniques which minimize temperature sensitivity and which allow unique phase variation control with a minimum drive power requirement are described. Peak and average power capability of the phase-shifters are discussed.

1. INTRODUCTION

Microwave phasors using ferrites or similar materials have until recently suffered from two main drawbacks:

(1) Temperature sensitivity, and

(2) Slow switching speeds with large switching power requirements.

During recent years, these problems have been largely overcome with the use of the same ferrite materials in tubular form, so that the ferrite can be magnetized circumferentially by a single wire passing through the center of the tube. Since many microwave ferrites exhibit a high degree of remanence, the ferrite can be switched to any state of remanent magnetization, and the RF phase information will remain without a continuous bias.

It is the purpose of this paper to discuss some of the problems associated with this type of phasor, including peak power limitations, temperature stability, and switching energy requirements.

This type of phasor permits the realization of a new class of digitally controlled devices as well as the usual phased array element by virtue of its latching behavior. In conjunction with hybrids and tees, the phasors can be varied to yield digital phase measuring circuits; digital phase discrimination devices; and amplitude and phase compensating networks. Drive circuitry which will allow very simple, but accurate control of the phasors, and which reduce the temperature sensitivity problem to one of a secondary nature are described.

* The research reported here was performed under subcontract to Lincoln Laboratory, a center for research operated by Massachusetts Institute of Technology, with joint support of the U.S. Army, Navy, and Air Force. For further information on these topics, the reader is referred to *Final Report* on "Development of Helical Phase-Shifters," Lincoln Laboratories, subcontract 250, Prime contract No. AF 19(628)-500.

2. THEORY OF OPERATION

Probably the most well-known latching phase shifter is the rectangular waveguide device first described by Levey and Silbur.[1] A sketch of this device is shown in Fig. 1. The ferrite tube can be square or round cross-section. A current pulse through the switching wire can magnetize the ferrite tube in either circumferential direction. The vertical walls of the ferrite tube are arranged to be in the position of circular polarization of the r.f. magnetic fields, and the interaction between the r.f. magnetic fields and

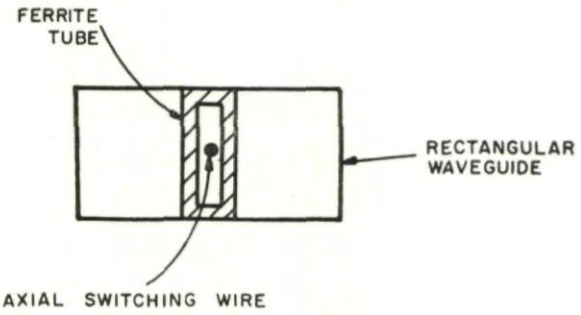


FIG. 1. *Waveguide latching phase-shifter*

the magnetic moments of the ferrite tube can be varied by simply changing the direction of the magnetization or by varying the relative number of moments oriented in either circumferential direction (partial magnetization). The effective r.f. permeability can, therefore, be varied and controllable phase-delay is obtained.

This waveguide device is practical in the upper end of the microwave frequency range where the physical dimensions, weight, and driving energy are not prohibitive.

At the lower microwave frequencies, a more compact device is the helical phase shifter.[2] A picture of this device is shown in Fig. 2. A helix is wrapped on a grooved cylinder of boron nitride, in the center of which is a ferrite tube. The helix is closely shielded to avoid dispersion, and to maximize the bandwidth of the device. In Fig. 3 is shown the resultant r.f. magnetic field due to the current flowing in two adjacent turns of the helix. At points A and B, the r.f. magnetic field lines due to the current in the helix turns are spatially orthogonal. The relative time phase of the r.f. fields due to the current in adjacent turns is a function of the helix diameter, and in particular if the helix diameter is chosen such that the circumferential distance per turn is an odd multiple of quarter wavelengths, the r.f. magnetic fields at points A and B of Fig. 3 will be orthogonal in time phase. The resultant r.f. magnetic field will then be circularly polarized relative to the direction into the paper. A qualitative curve showing the variation of the type and sense

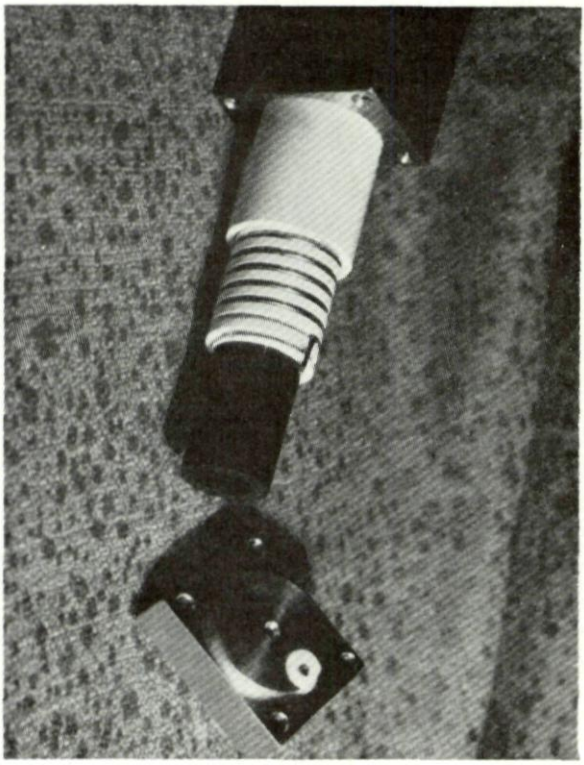


FIG. 2. *S-band helical phase shifter*

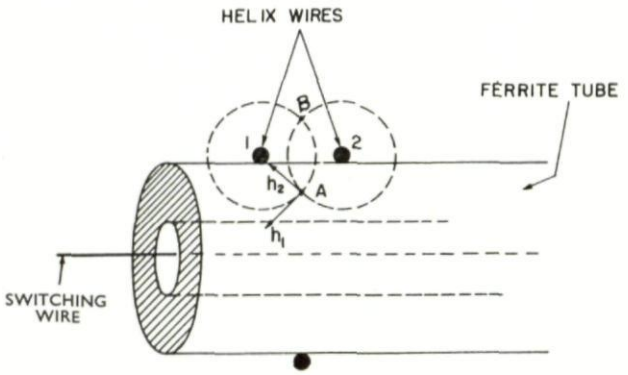


FIG. 3. *Resultant R.F. Magnetic field in ferrite wall*

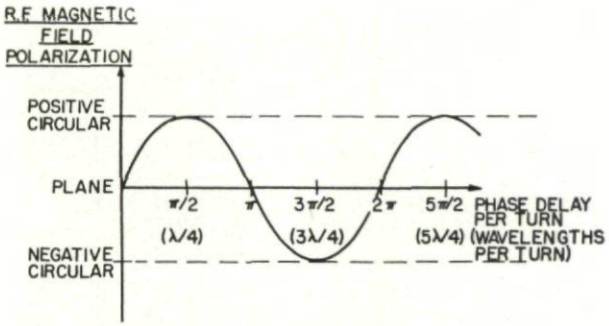


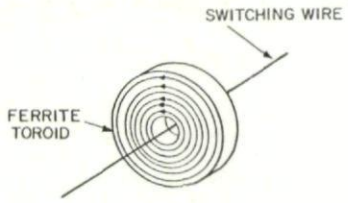
FIG. 4. *Polarization of resultant RF magnetic field in ferrite wall as a function of the phase delay per turn of the helix*

of the polarization of the resultant field at A and B as a function of wavelengths per turn of helix is shown in Fig. 4.

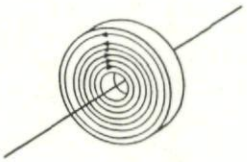
It is clear from this polarization curve, that the differential phase characteristic of the helical phase-shifter has a somewhat sinusoidal frequency dependence. The fractional bandwidth is a maximum if the helix is $\lambda/4$ per turn at the center frequency. For mechanical reasons, the phase shifters are usually designed with 3/4 wavelength per turn.

3. PARTIALLY MAGNETIZED FERRITE

Ideally, we can think of the ferrite toroid as being composed of a very large number of infinitely thin concentric tubes (Fig. 5). If these tubes are



(a) MAGNETIZED TO SATURATION



(b) PARTIALLY MAGNETIZED

FIG. 5. *Magnetization mechanism*

all aligned in one circumferential direction, the ferrite is magnetized to saturation. The ferrite is only partially magnetized if some of the tubes are oriented in the opposite direction.

If the r.f. magnetic field is circularly polarized at all points in the ferrite wall, the moments oriented in one direction will interact and yield the permeability μ_+, while the moments oriented in the opposite direction yield

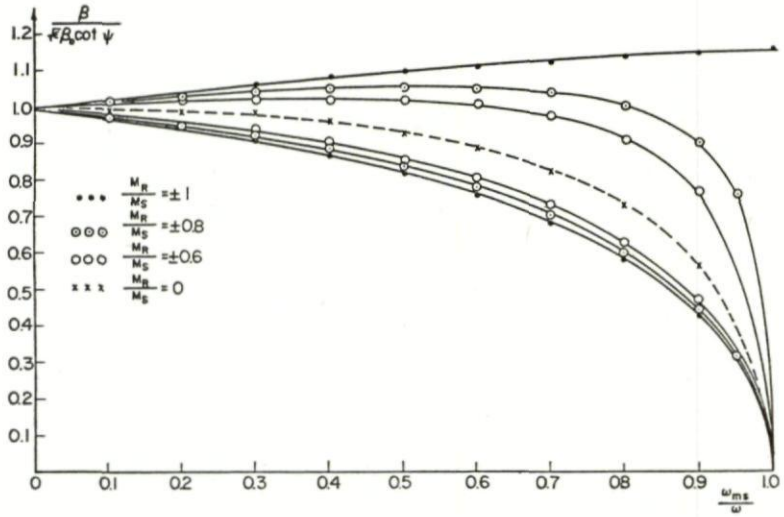


FIG. 6. *Theoretical propagation characteristic as function of* (ω_{ms}/ω) *with* (M_R/M_S) *as parameter*

the permeability μ_-. The resultant ferrite permeability is a weighted parallel combination of μ_+ and μ_-.

Using this idealized model, the propagation characteristic of the helical phase-shifter can be obtained as a function of remanent and saturation magnetization. This characteristic is shown in Fig. 6. for a particular ferrite loading factor in the helical phase-shifter.

It should be noted in Fig. 6 that in certain states of magnetization the device is quite insensitive to variations in saturation magnetization provided the ratio M_R/M_S does not vary. It turns out, that for most of the microwave ferrites this condition is met if the ferrite is driven with a field much greater than its coercive field. The maximum ratio M_R/M_S for many ferrites or garnets is in the range 0.5 to 0.8. We can, therefore, arrange to have a temperature stable reference phase simply by overdriving the device in one direction.

In Fig. 7 is shown the behavior of the phase constant with the remanent magnetization as parameter. It is readily seen that with a temperature independent change in magnetization, the differential phase shift between

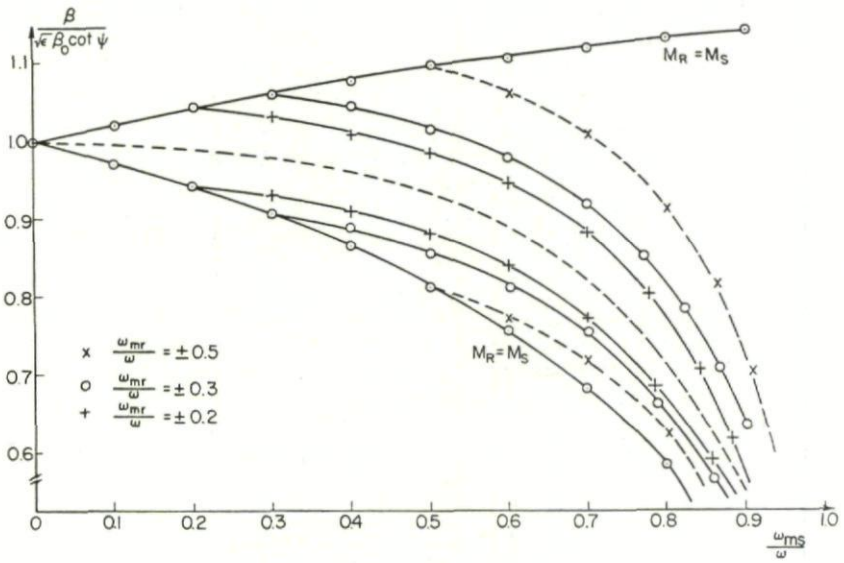


FIG. 7. *Theoretical propagation characteristic as a function of* (ω_{ms}/ω) *with* M_R *as a parameter*

any two states is independent of variations in the saturation magnetization of the ferrite.

Such a temperature independent change in net remanent magnetization can be achieved with the use of flux transfer techniques.

4. DRIVE CONTROL CIRCUITRY

4.1 *Flux Transfer*

The basic idea of flux transfer is shown in Fig. 8. The saturable drive core is switched always to saturation, producing a constant volt-time product across the phase-shifter core. The flux, and therefore the remanent

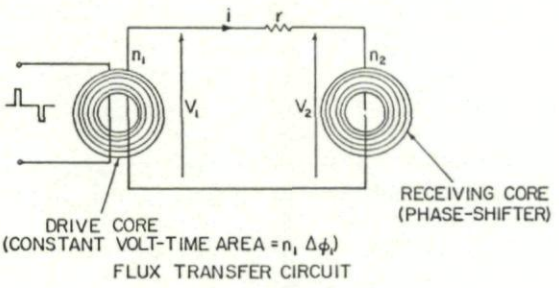


FIG. 8. *Flux transfer circuit*

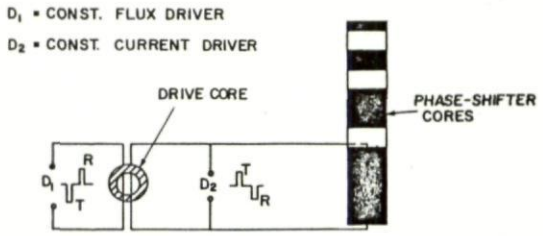


FIG. 9. *Cascaded binary phasor*

magnetization, of the phase shifter core is changed by an amount depending only on the drive core. Since the drive core need not be a microwave ferrite, it can be chosen for its temperature stability and squareness properties. If the drive core is not temperature stable, it can be placed inside a thermostatically controlled oven.

4.2 *Phase Control Techniques*

Different types of phase-control techniques are possible with the latching phase-shifter. The type chosen depends on the desired phase-shifter function. For instance, in phased array applications where random beam positions are required, a binary drive control allows the phasor to be set at any value of phase in discrete steps which are multiples of the smallest bit.

In Fig. 9, the binary phase increments are obtained simply by cascading the required number and lengths of ferrite tubes which are spaced apart by

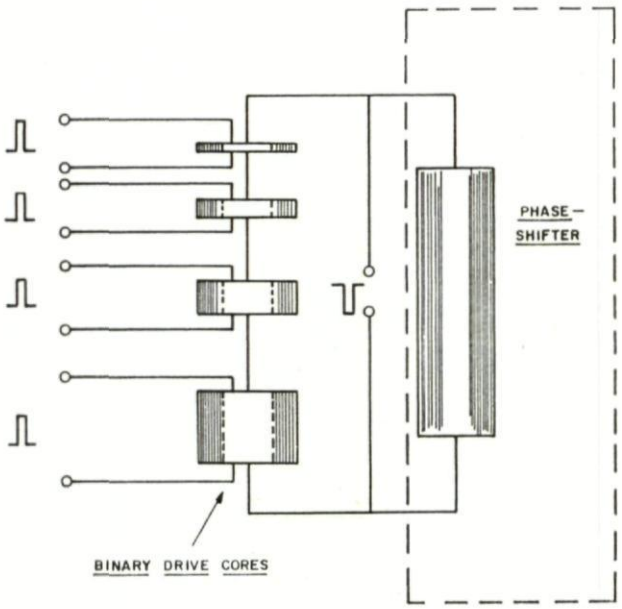
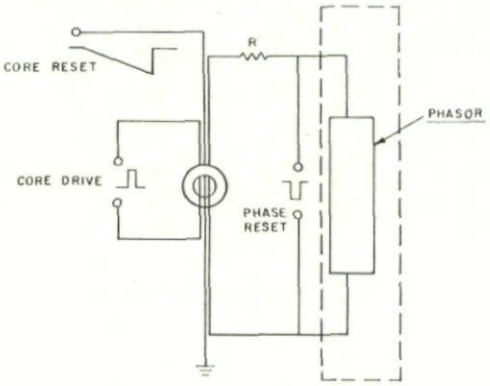


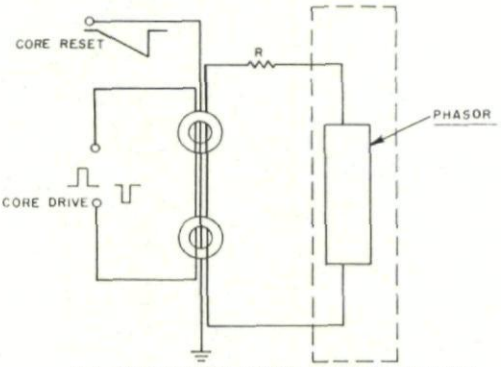
FIG. 10. *Partial switching in binary bits*

means of non-magnetic dielectrics. The drive cores are used simply to obtain temperature stability and a means for final adjustment of the bit-sizes.

In Fig. 10, a single ferrite tube is used in the phase shifter, and the device is partially switched to the desired phase setting by means of binary drive cores,



(a) UNIDIRECTIONAL PHASE STEPPING CONTROL



(b) BIDIRECTIONAL PHASE STEPPING CONTROL

FIG. 11. *Uniform phase stepping control*

as shown. Again in this case, the drive cores provide the temperature stability and the bit-sizes can be adjusted externally to the microwave circuit.

Other types of control which do not allow random phase setting, but which simply increase or decrease the phase delay in equal increments are illustrated in Fig. 11. In the top example, the phase delay is simply decreased in equal steps each time the drive core is pulsed. The phasor is reset to the temperature stable state before each pulse train.

In the second example in Fig. 11, the phasor can be stepped up or down in equal increments depending on which drive core is pulsed. This device will not be temperature stable unless the phasor is reset periodically.

5. SWITCHING ENERGY REQUIREMENTS

Since these devices will in general be operated by low level computer signals, an amplifying switch is required between the computer and the drive cores. The complexity and cost of this switch will depend on the required switching time, the maximum repetition rate, and the amount of flux to be switched.

A switching figure of merit can be defined for a given microwave circuit as the phase change obtained per unit of magnetic flux. This unit of magnetic flux can be conveniently chosen as the microweber, or one volt-microsecond ($V\mu s$).

With the helical phasor, the switching figure of merit is of the order of $25°/V\mu s$. For the waveguide device, the number is of the order of $5°/V\mu s$.

The current required during the switching period depends on the coercive force of the ferrite, and the outer diameter of the ferrite tube. For many of the ferrites used, the coercive force is of the order of one oersted. A ferrite tube 1 cm diameter will, therefore, require 2.5 amperes drive current.

Thus, to switch a 360° phasor in one microsecond requires a switch which delivers greater than 2.5 amperes at 15 volts for the helical device (37.5 microjoules) and 75 volts for the waveguide device (190 microjoules).

Such a switch can be realized with the use of state-of-the-art high power transistors or silicon-controlled rectifiers (S.C.R.). The S.C.R. suffers from the disadvantage that it takes longer than 10 μs to turn off. However, the S.C.R. can be operated directly from computer signal levels, and such a switch requires no standby biasing power.

The transistor switch, on the other hand, is capable of higher repetition rates (500 kc/s), but requires amplifying stages between the computer and the power transistor. The biasing power required for the transistor switches in a four-bit phasor can be of the order of one or two watts.

6. POWER HANDLING CAPABILITY

6.1 *Peak Power*

The main limitation to the peak power capability of the devices is the ferrite non-linearity at high r.f. field levels.

The threshold field strength appears to the relatively independent of the state of magnetization so long as the ferrrite is not saturated, and depends on the normalized saturation magnetization ($4\pi M_S/\omega$) of the material.

A possible mechanism for the dominant non-linear effect in the useful frequency range of the ferrites is simple subharmonic generation caused by

parallel pumping in the domains. Such a model has been analysed, and the results are shown in Fig. 12. Also shown are some experimental data.

In Fig. 13 is shown the actual threshold power level obtained in the helical phase-shifter.

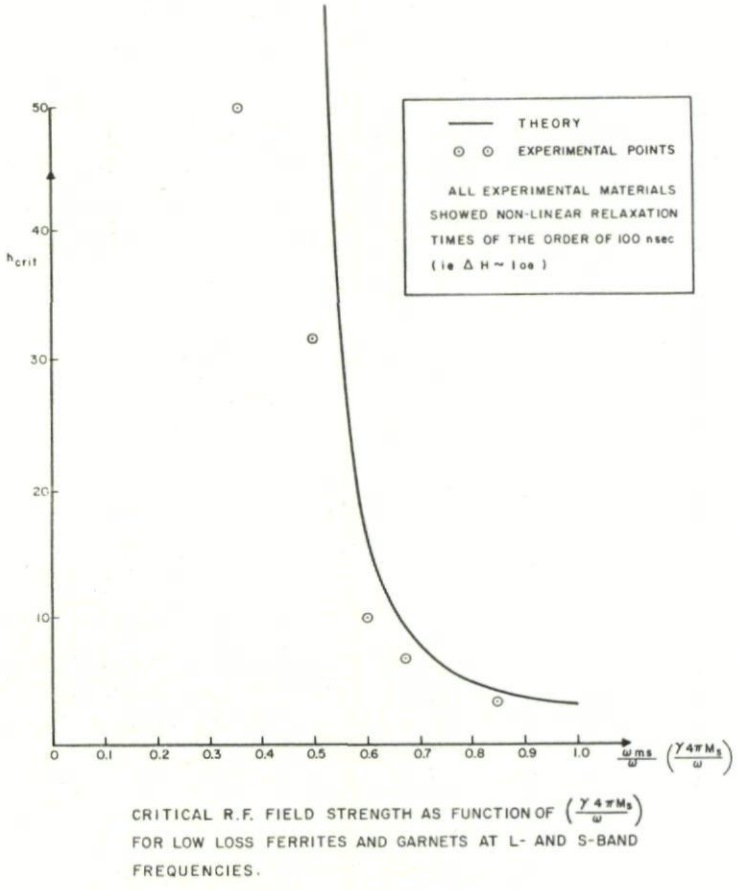


FIG. 12. *Critical r.f. field strength as function of* $(\gamma 4\pi M_S/\omega)$ *for low loss ferrites and garnets at L- and S-band frequencies*

6.2 *Average Power*

The use of boron nitride as a dielectric medium between the ferrite and the outer shield gives a thermal path with conductivity of the order of one tenth that of copper. In an *S*-band 360° phasor with 1 db insertion loss, an average power of 200 watts out of the device resulted in 50 watts dissipated power. The temperature difference between the ferrite and the outer shield was 10 °C under these conditions.

Using the saturable core driver techniques, the saturation magnetization

of the microwave ferrite can be allowed to vary by approximately 40% with no appreciable deterioration in the phase characteristics. The allowable temperature range, therefore, depends on the Curie temperature of the ferrite, and for uncompensated materials is approximately 40% of the Curie temperature reckoned from 0 °C.

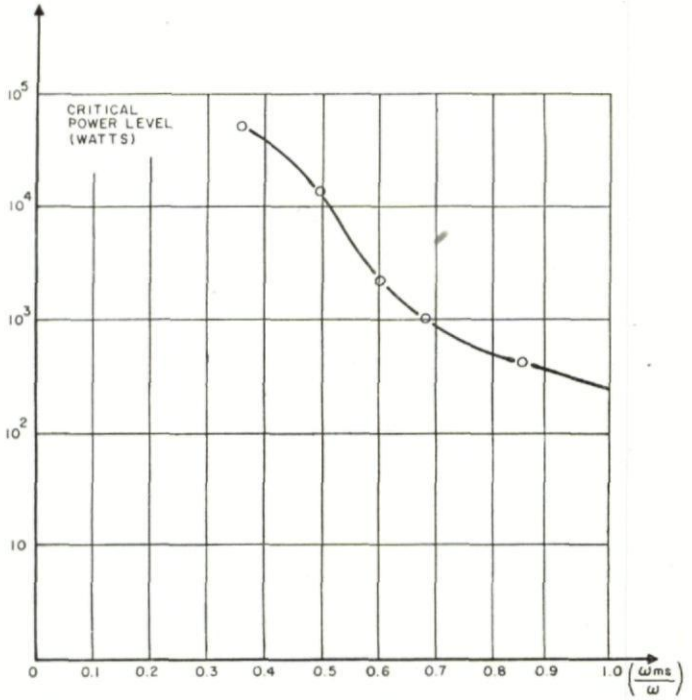


FIG. 13. *Experimental critical power level of Helical phase-shifter at L-and S-band frequencies using low loss ferrite and garnet materials*

DISCUSSION

E. R. CASHEN: Does Dr. Hair know of any ferrite ring configuration giving a reciprocal phase shift?

H. A. HAIR: There is no reciprocal type that is capable of 1 μs switching.

B. CUTLER: The paper quotes 1 db loss per 360° phase shift at *S*-band. What performance is possible at *X*-band?

H. A. HAIR: For low power devices in waveguides 1 db loss would be associated with 1000° phase change. At the higher powers a ratio of 360°/db is possible.

REFERENCES

1. LEVEY, L. and SILBER, L. M., A fast switching *X*-band circulator using ferrite toroids (Wescon Convention 1960).
2. Final Report on Development of Helical Phase-Shifters, Lincoln Lab. Subcontract 250, Prime contract AF19(628)—500.

CHAPTER 26

DISPERSORS FOR PULSE COMPRESSION SYSTEMS

P. S. Brandon, O. E. Keall, and W. S. Mortley
The Marconi Co. Ltd., Chelmsford, England

This paper deals with two types of apparatus for producing the dispersion required in a pulse compression system. The first type is composed of lumped electric circuits and has reached an advanced level of development; the second makes use of ultrasonic techniques and is in the experimental stage.

The design of the electric circuit dispersor is dealt with in Part 1, *and the realisation of a typical design in Part* 2. *Part* 3 *deals with the ultrasonic dispersor.*

PART 1

DESIGN OF LUMPED CONSTANT DISPERSIVE NETWORKS*

P. S. Brandon

Figure 1 shows the variation of delay, τ, with frequency, f, required for a pulse compression receiver to give maximum signal against a random background—noise, uniform clutter, etc. It is assumed that the transmitter is frequency modulated linearly with time. A weighting filter is required to give an optimum amplitude shape for radar pulse compression, and will be discussed later.

Figure 2 shows the building brick for an all-pass constant-R network suitable for building networks with dispersive characteristics. It should be noticed that it has a τ/f characteristic of power resonance shape, of width Δf about the tuning frequency f_s. An effective Q, known as Q_e, can be associated with this network; it is approximately $f_s/\Delta f$. Note the relations shown on this figure, particularly No. 3. Figure 3 shows how the shape of the curve of τ against f changes with Q_e value. Figure 4 shows how the component values vary with f_s and Q_e value. L is proportional to R, and C inversely proportional to R. By choosing $R = 70\ \Omega$ the desirable ranges of

* Based on a paper presented at a meeting of the Radar and Navigational Aids Group of the Institute of Electronic and Radio Engineers, and published in the Institution's journal "*Radio and Electronics Engineer.*"

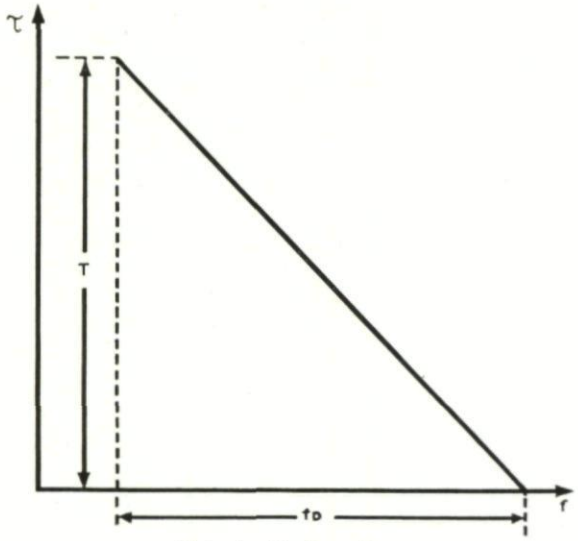
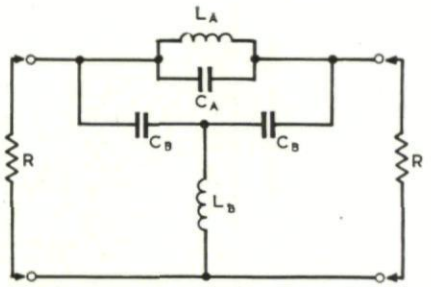


FIG. 1. *Ideal requirement.*



1. $C_B = \frac{1}{R\omega_s Q_e}$ WHERE $Q_e \approx \frac{f_s}{\Delta f}$
2. L_B TUNES WITH $2C_B$ TO f_s
3. $L_A = \frac{L_B}{Q_e^2}$
4. C_A IN PARALLEL WITH $\frac{1}{2}$ C_B TUNES WITH L_A TO f_s

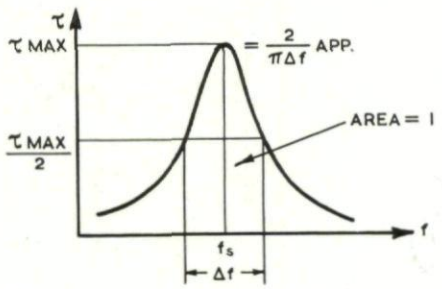


FIG. 2. *Delay characteristics.*

values of L and C (dashed lines) almost coincide. From this it can be seen that Q_e should be less than 10 for higher frequencies and less than 5 below 10 mc/s or so.

Figure 5 shows the principles upon which a dispersive characteristic can be built. The top example is composed of equally spaced sections with

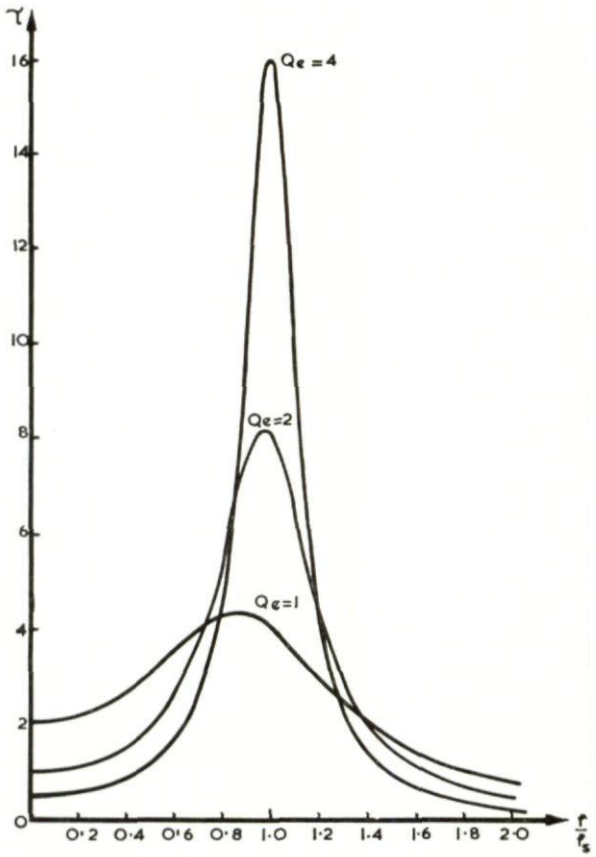


FIG. 3. *Effect of various Q_e values.*

narrower bandwidth at the high-delay end of the band. The bottom example is made up of equal bandwidth circuits packed more densely at the high-delay end.

Going into this more deeply we must first observe that the characteristic of a single all-pass section can be represented as the "mirror image" sum of two symmetrical shapes, as shown in Fig. 6. Figure 7 illustrates how a constant delay can be built up from an infinite series of these symmetrical shapes. They are spaced at equal intervals f_σ and give a nearly constant

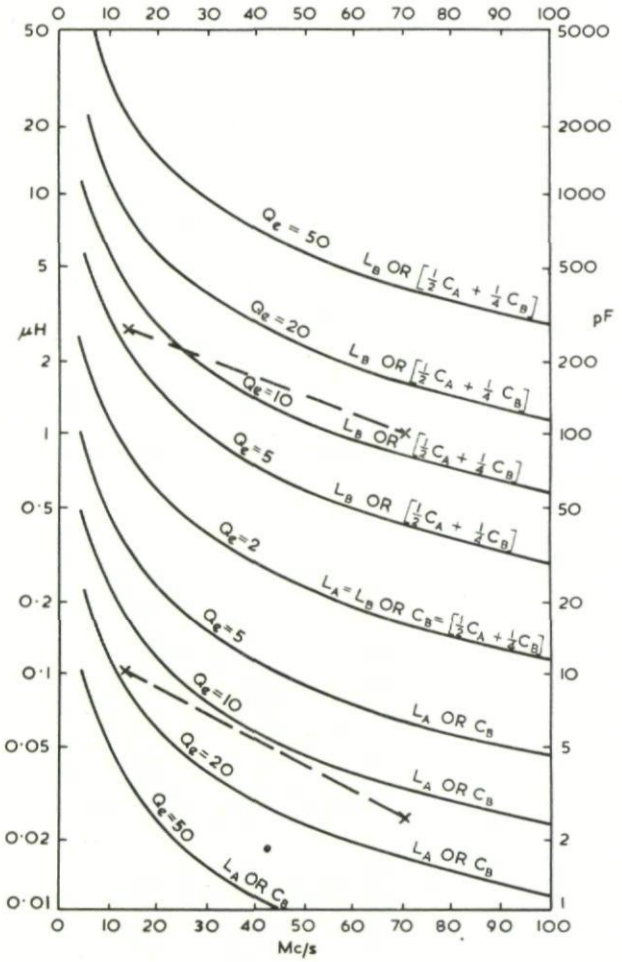


FIG. 4. *Component values for* $R = 70\ \Omega$.

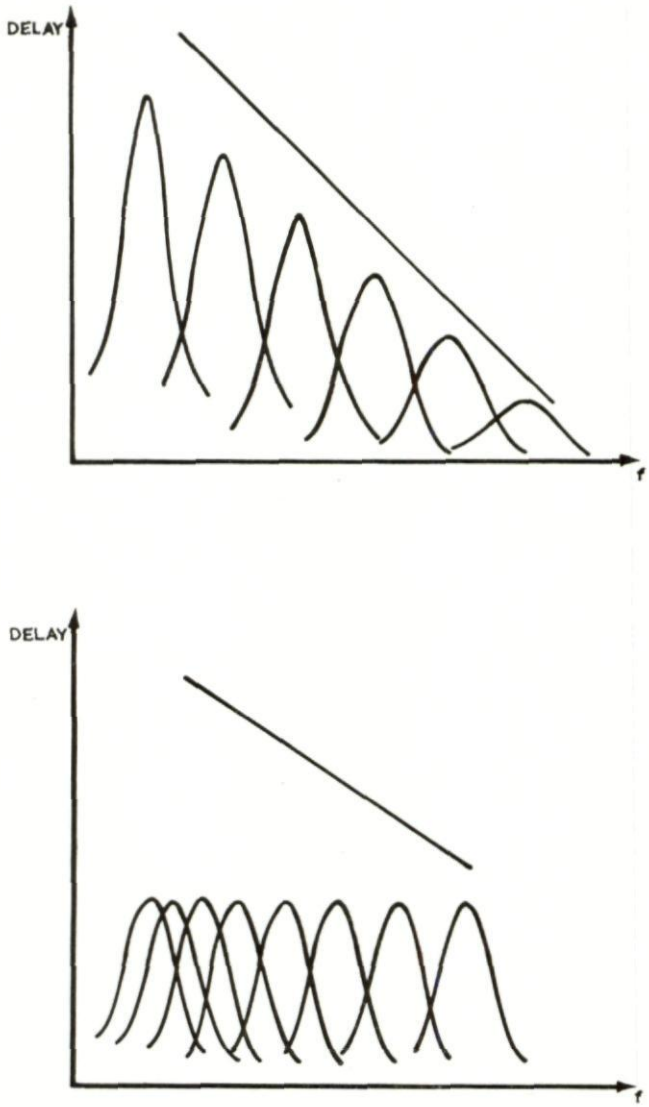


FIG. 5. *Methods of combining all pass sections to give dispersion.*

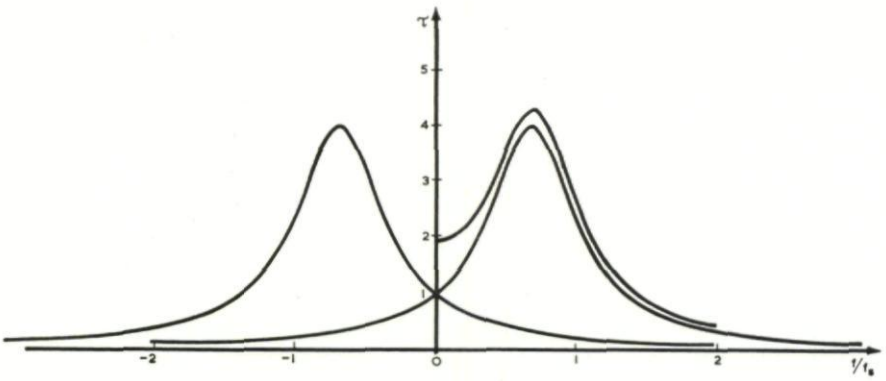


FIG. 6. *Characteristics of a single all pass section considered as the mirror image sum of the two symmetrical shapes.*

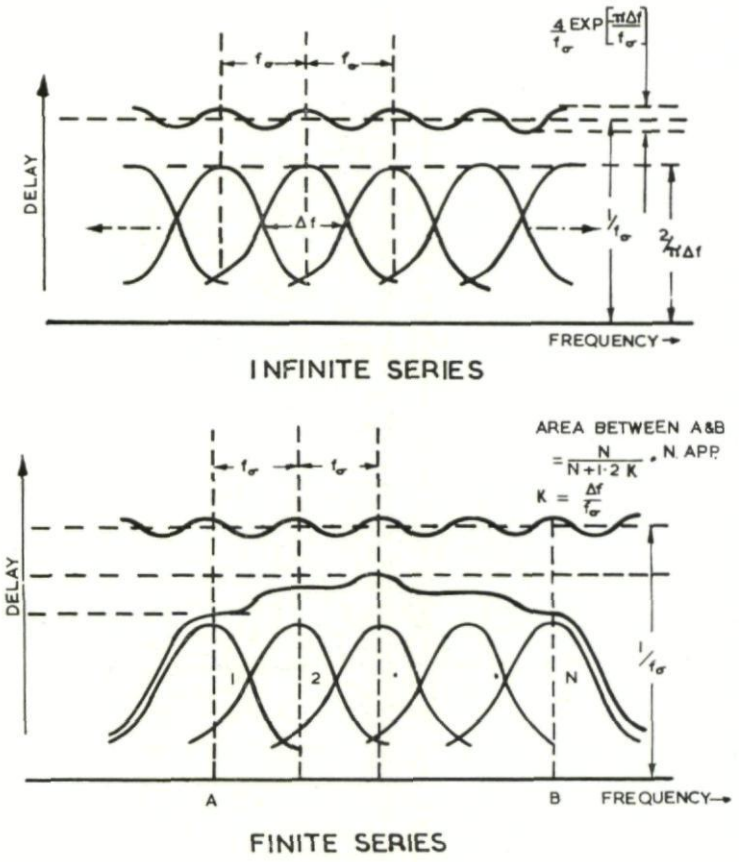


FIG. 7. *Combinations for constant delay.*

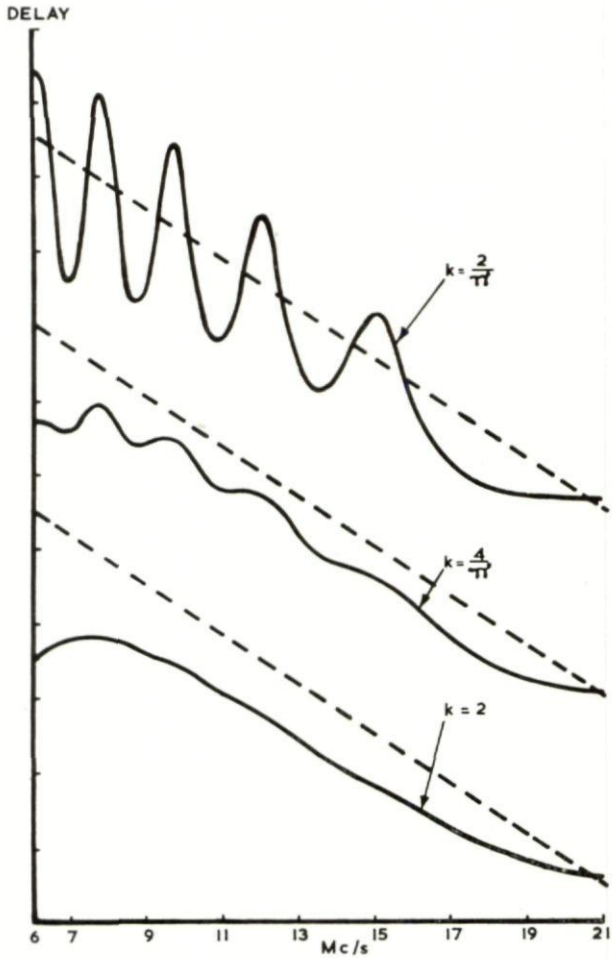


FIG. 8. *Effect of K value on ripple.*

delay $1/f_\sigma$. The ripple is a function of $\Delta f/f_\sigma = k$, say. Using the simple rule of making f_σ inversely proportional to the required delay and keeping k constant gives the approximations to the dashed lines shown in Fig. 8. For $k < 2$ a ripple appears at "section spacing"; for $k > 2$ the ripple disappears, but the approximation is less good.

Figure 9 shows the delay characteristics of a series of designs covering 11–16 mc/s, each design being made up of sections with the same Q_e. Keeping Q_e constant was found to be a better design procedure than keeping k constant, as component values were more under control. It can be seen from this figure that the lower Q_e the greater is the "overspill" outside the band and consequently the greater the number of sections required for the

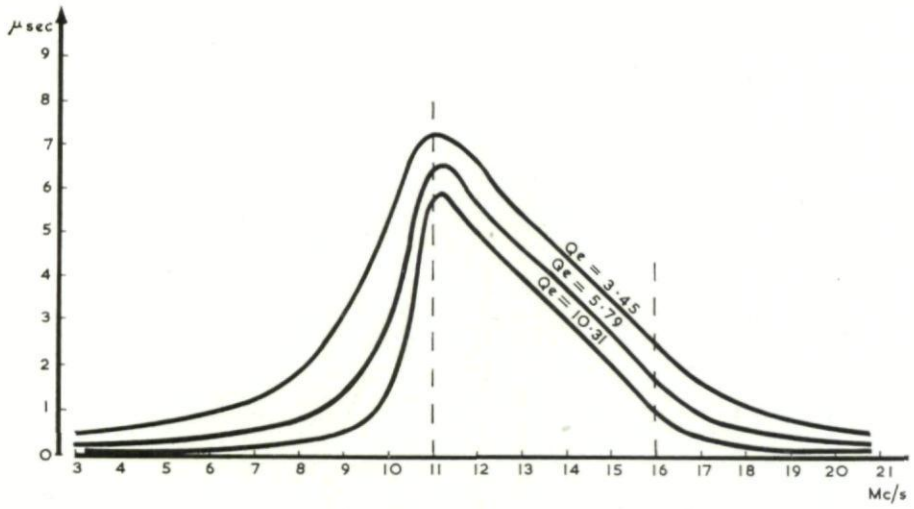


FIG. 9. *The effect of Q_e on overspill.*

same differential delay T. As is to be expected, a compromise between awkwardness of components and number of sections must be reached.

Figure 10 illustrates the first step in a practical design method. The network can be made up of G identical groups, each of n_a sections; it is found that $G = f\tau/kQ_e$, all quantities having their mid-band values. The range of delay to be provided by each group is T/G; n_a would be equal to

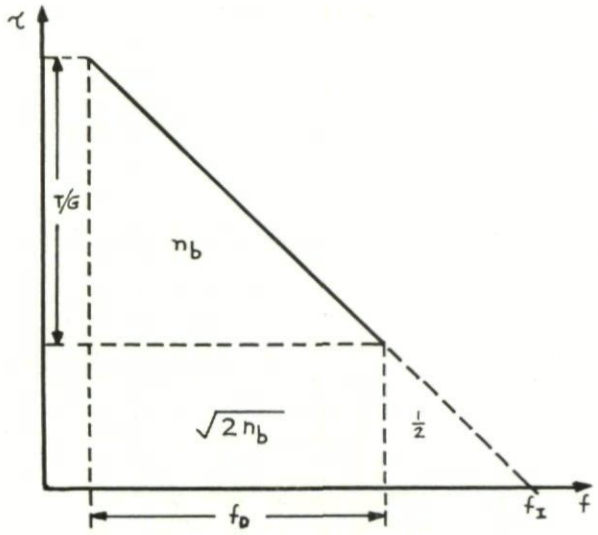


FIG. 10. *First step in a practical design procedure.*

n_b in the figure if there were no overspill effect. At the high-frequency end, overspill can be allowed for by adding the triangle of half unit area, which means raising the specification on the "pedestal" of area $\sqrt{2n_b}$. To cover the overspill areas completely it is found that an additional $1.2k$ sections are required. Thus $n_a = n_b + \sqrt{2n_b} + 1.2k$. The n_a sections are arranged as

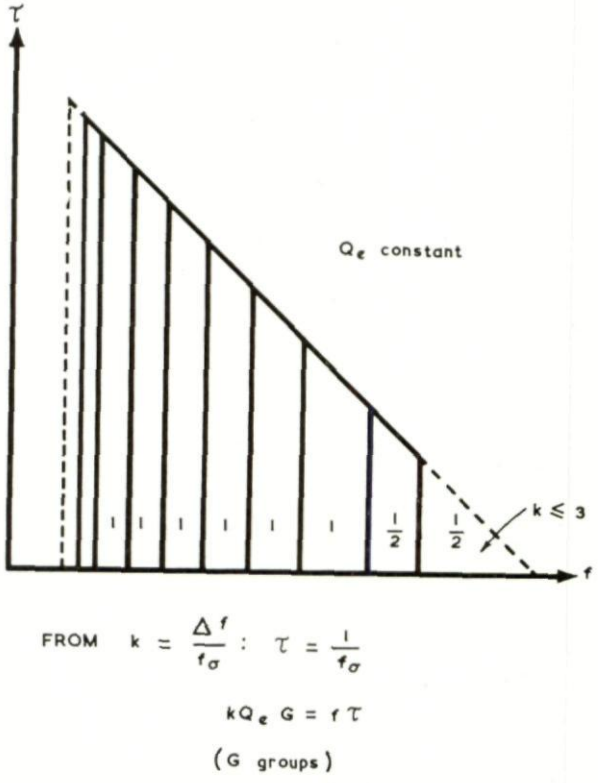


FIG. 11. *The arrangement of sections.*

in Fig. 11, the first at half a unit area from the high-frequency edge of the band (one unit area from the end of the modified characteristic) and the rest at successive unit areas. Such a design has the maximum number of identical groups.

This method can be used with $k = 3$ to 4 and $Q_e = 4$ to 5 to give a good practical design, but the approximation to the required characteristic is rather poor. Figure 12 shows how this approximation (lowest curve) can be improved with the help of a computer. The original specification (dashed line) is distorted in proportion to the departure of the original approximation from it. Using this new specification, and spacing sections at unit area points as before, a new design is obtained giving a closer approximation to the original specification (top curve). This process can be repeated indefinitely.

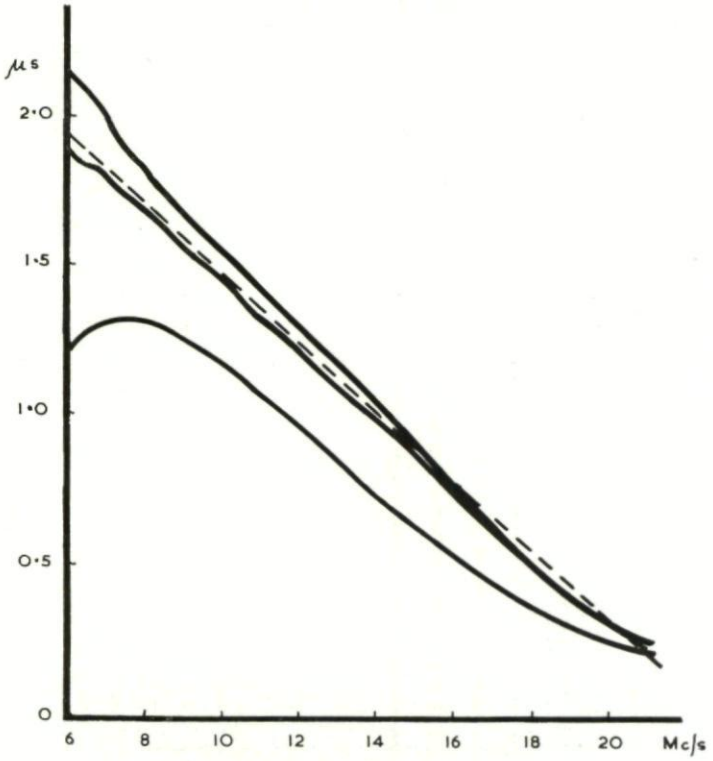


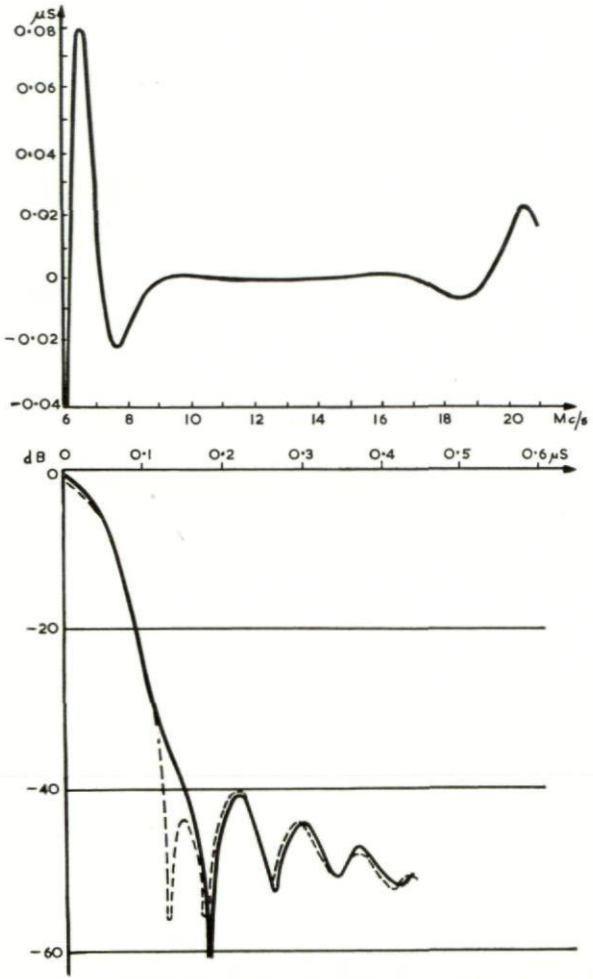
FIG. 12 *Approximations to specified characteristic:*

Dashed curve	*Specified characteristics*
Lower full curve	*First approximation*
Top full curve	*Second approximation*
Middle full curve	*Third approximation*

After one further predistortion the middle solid curve in Fig. 12 is obtained; by repeating the process until there is no further improvement, the error remaining is typically as shown at the top of Fig. 13. This curve refers to a design with $T = 10$ μs, $f_D = 15$ mc/s (between 6 and 21 mc/s), having 6 groups of 23 sections each. The pulse sidelobes introduced by such an error are shown at the bottom of the figure; to find the total effect it is necessary to add the side lobes due to the edges of the signal and of the dispersive characteristic, shown in Fig. 14.

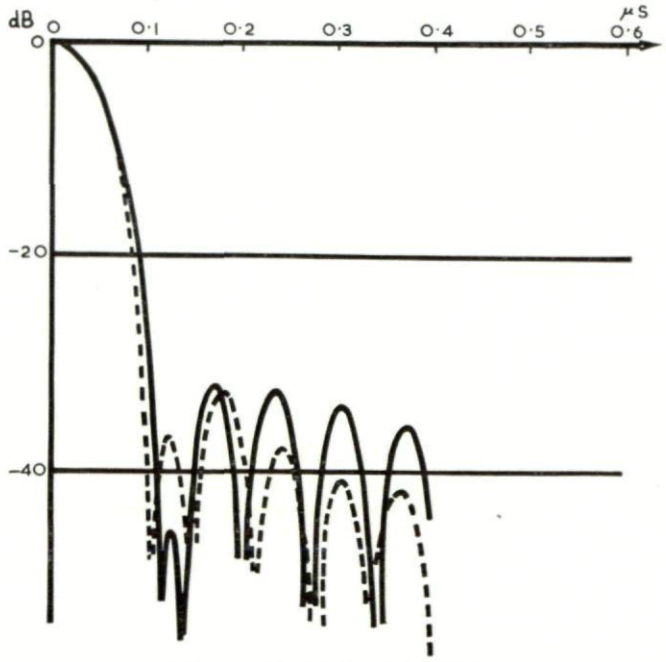
Figure 15 shows the improvement obtained by taking the design of Fig. 13 and slightly modifying the Q_e values (range 3.88 to 4.123 instead of 4 for all sections). This method is very useful, particularly in narrower band designs.

Figures 16 and 17 indicate how the effect of parasitic components can be allowed for. Figure 16 shows the method of allowing for the lead inductance of C_A and the stray capacity of L_B. Losses in components destroy the all-pass characteristic of a network and have to be compensated for by a suitable



FULL LINE, TRAILING EDGE;
BROKEN LINE, LEADING EDGE.

FIG. 13. *The characteristics of a network with* 6 *groups each of* 23 *sections* ($T = 10$ μs *and* $f_D = 15$ mc/s).



FULL LINE, TRAILING EDGE;
BROKEN LINE, LEADING EDGE.

FIG. 14. *The characteristics of network showing total side lobes.*

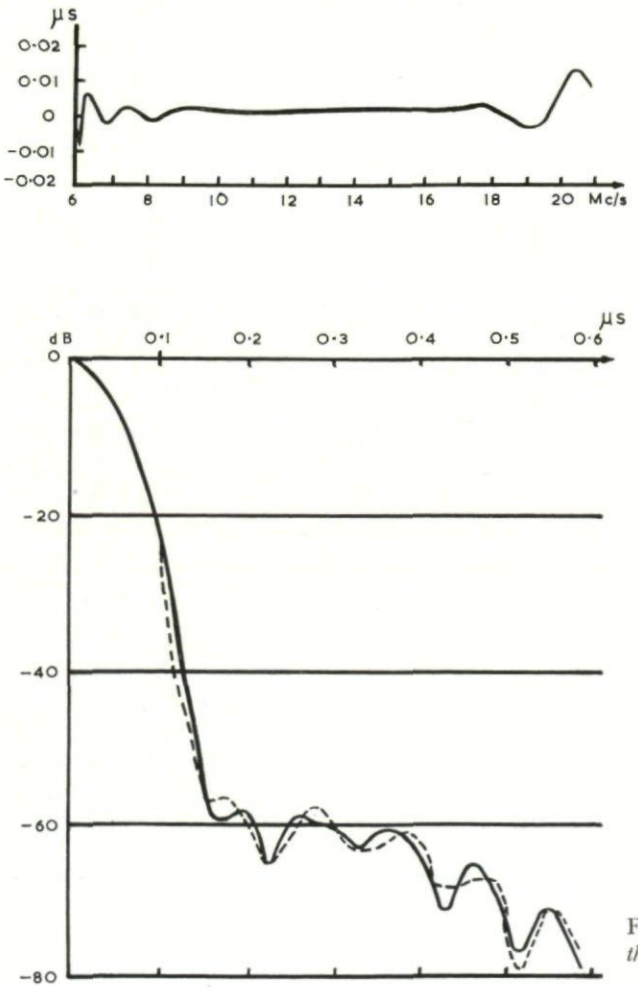


FIG. 15. *Network of Fig.* 13 *improved by modifying the* Q_e *values (range* 3.88 *to* 4.123 *instead of* Q_e = 4.00 *for all sections).*

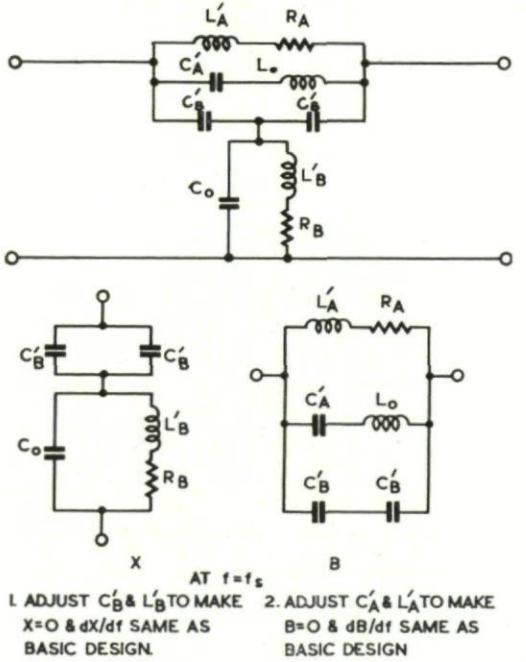


FIG. 16. *The correction for lead inductance in* C_A *and stray capacity in* L_B.

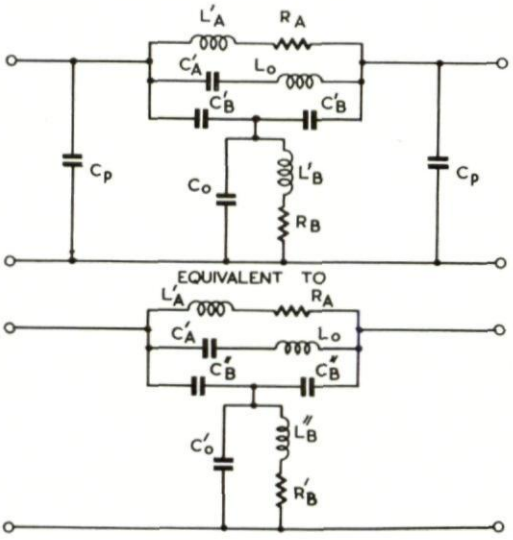


FIG. 17. *The correction for stray input and output capacities.*

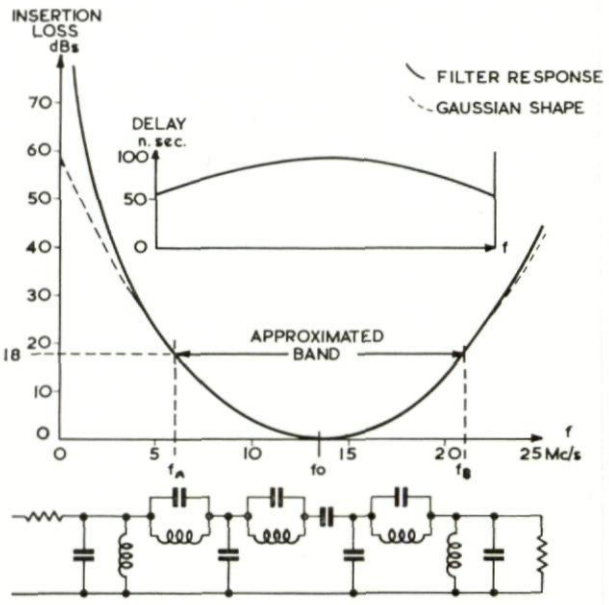


FIG. 18. *Amplitude weighing filter for side lobe control.*

equaliser. Particular care must be taken to compensate for the skirts of the "band-stop" characteristic introduced by the losses. Figure 17 shows how stray capacities across input and output can be allowed for by the use of an equivalent circuit in the form of Fig. 16. When properly allowed for in the process of alignment, the effect of parasites and losses can in fact be made negligible. The main practical problem is the alignment of the sections with sufficient accuracy.

Figure 18 shows the kind of amplitude-weighting filter used to give an overall 30–34 db range side lobe performance. This has a characteristic shaped like a Gaussian error curve, giving 18 db attenuation at the edges of the band, i.e. in this case 6 and 21 mc/s. The circuit is designed by standard methods of approximation. It must be remembered that such a network has a non-constant delay characteristic, which may be troublesome. This can be improved on occasion by the use of two networks in cascade, for example, each giving 9 db attenuation at the edges of the band. This is particularly true for narrower bandwidths than that shown.

PART 2

REALISATION OF A DISPERSIVE NETWORK SYSTEM

O. E. KEALL

INTRODUCTION

A dispersive network system will usually consist of

(1) a network or other device in which the group delay of a signal passing through it varies with its frequency,

(2) an equalising amplifier or network to compensate for the variations of attenuation that usually occur within the dispersive network over the operating band, and finally

(3) a weighting amplifier or filter the purpose of which is to control the level of the side lobes associated with the main lobe or pulse at the output of the system when the appropriate transient signal matched to the network characteristics is fed into it: without such weighting the side lobes or pulses have a $(\sin \theta)/\theta$ form, the first side lobe being but 14 db below the main pulse and thus liable to be mis-interpreted in a radar system as a target.

A dispersive system built to provide a variation of delay of 10 μs over a 14 mc/s bandwidth operates between 70 Ω impedances at a centre frequency of 13.5 mc/s. The advantage of a relatively low centre frequency is that the number of all-pass sections required to achieve the desired performance is reduced while component values are large relative to stray and parasitic

elements but not so large that the self-resonant frequencies of components are too close to the operating band.

1. *The Dispersive Network*

The network was constructed of six similar groups each providing 2 μs delay at 6 mc/s decreasing linearly to 0.333 μs at 21 mc/s, a variation of 1.67 μs delay over the band. 23 all-pass bridged-T sections of $Q_e = 4$ were used in each group, the sections being staggered asymmetrically over the frequency band as described in Part 1. The network thus consisted of 138 sections in cascade: ultimately a single compensating section was added to improve side lobe performance.

2. *Equalising Network*

It has not so far been found practical to estimate in advance the insertion-loss characteristic of a completed network; the equalising network is thus computed from the measured insertion-loss characteristic of a prototype group of sections from which the final network is to be assembled (or from the completed network itself). Ideally, the equalising network should, when associated with the dispersive network, provide constant attenuation within the operating band and high attenuation outside that band while introducting small but constant delay within the working band. The rapid change in attenuation of a dispersive network from a high to a low value as frequency is reduced below the lower band-edge frequency has presented some difficulty and it is preferred as a result not to work with overall band widths less than a fifth of the centre frequency, but the difficulty could be overcome by associating an equaliser with each group of sections. Due to the fact that equalising networks generally present wide variations of impedance within the working band, they cannot be connected directly to a dispersive network, which must be correctly terminated in order to avoid internal reflections. It is therefore convenient to use valves or transistors for isolation of the network as well as to provide gain to overcome system loss: in this connection it is customary to specify gain so as to provide constant peak pulse level at input and output of the system or alternatively to provide unity c.w. gain at the centre frequency, in which case the output pulse level will exceed the input pulse level by, nominally, the square root of the compression ratio Tf_D.

3. *Weighting Filter*

For control of output side lobe level a weighting filter reducing first side lobes to -25 db and others to -30 db or more is regarded as operationally acceptable; this performance may be obtained by means of a Gaussian shaped frequency response characteristic attenuating the band-edge signal frequencies 18 db relative to the centre frequency; since the response shape is Gaussian the phase characteristic of the filter is smooth and relatively linear and its effect on the phase characteristic of the network is minimal. As in the case of equalising networks, the change in impedance of a weighting filter over the working band causes excessive reflections in a dispersive

network and valves or transistors are used as isolating devices or to provide part of the gain required to meet the specification.

4. *Matching*

The importance of maintaining a good match not only at the network terminals but throughout the remainder of the system, both transmitting and receiving, must be stressed, as if this be not done, signals may be generated by reflection and be resolved by the network as true signals or side lobes, and performance will inevitably be degraded. The situation is more critical than in normal long-phase systems where reflections occurring within the pulse length cause small changes of amplitude and phase of duration short compared with the pulse length; in a frequency-swept system although a long pulse is used, the discrimination of the system is that of the short compressed pulse, and reflections that would be masked in a c.w. long-pulse system then become obvious and may set a limit to the amplitude discrimination of the system as a whole. There is thus no point in decreasing the side lobe level of a network by weighting much beyond the point at which the system itself begins to fail in discrimination.

5. *Final Adjustments*

It generally happens that, after the three major components of a system are associated with one another and measurements are completed, defects representing departures from a desired linear delay characteristic are apparent and one is faced with the problem of taking corrective measures. One will have a measured phase characteristic from which also the side lobe levels may be computed; a more practical demonstration of performance may be required as a condition of acceptance by a customer when the linearity of both the test source used to produce frequency swept pulses as well as the dispersive network may be in question. Although it is believed that both test source and network possess a high degree of linearity independently a much more rapid means of indicating this than is available by the point-by-point measurement methods used is becoming a necessity.

However, if one assumes the network to be at fault one can take corrective measures as follows:

(a) The characteristics of individual sections may be modified either by changing frequency slightly or by modifying Q_e. Such modifications have occasionally been carried out in the prototype stage (but are not recommended during production) to correct for peaks or dips occupying say 1/5 to 1/10 the total bandwidth; more rapid changes have little effect on side lobes whilst slower ones (less than 3 cycles within the band) are best dealt with using (b).

(b) The network is not altered in any way but a compensating section is added, the centre frequency and Q_e of which are adjusted to provide the additional delay needed to improve linearity. This method has proved relatively successful and 3 to 6 db reduction in side lobe level has been achieved by the addition of one section.

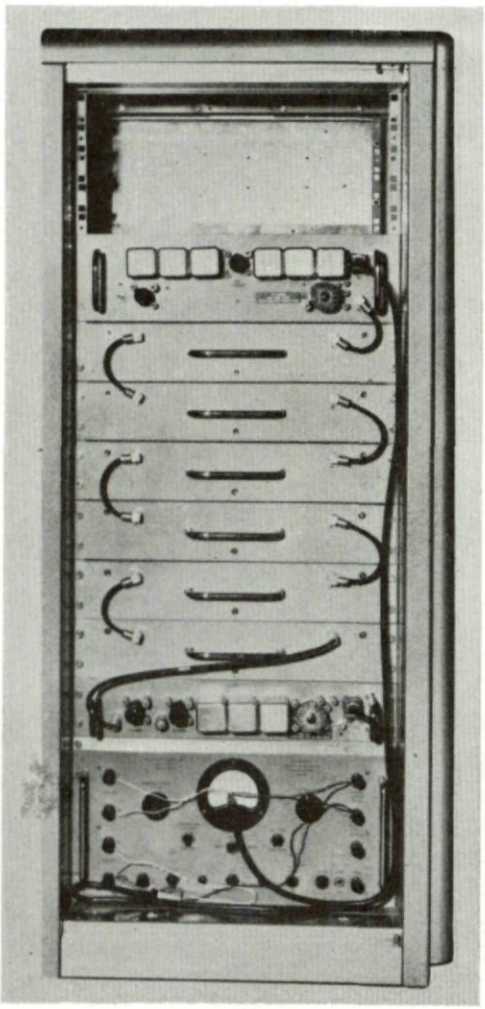


FIG. 19. *Complete dispersive system.*

6. *The Equipment and Its Performance*

A view of the complete dispersive system is shown in Fig. 19. The equipment was designed for ground station use and is therefore mounted in a 19″ rack, no special effort having been made to reduce the size. At the top is the weighting filter, then follow the six similar groups each of 23 bridged-T sections, whilst at the bottom is the equalising amplifier. The order in which the constituents are connected is not significant except that with some arrangements low noise factor and high dynamic range are more readily achieved. A view of one of the groups of sections with cover removed is given in Fig. 20. whilst details of a bridged-T circuit assembly are seen in Fig. 21.

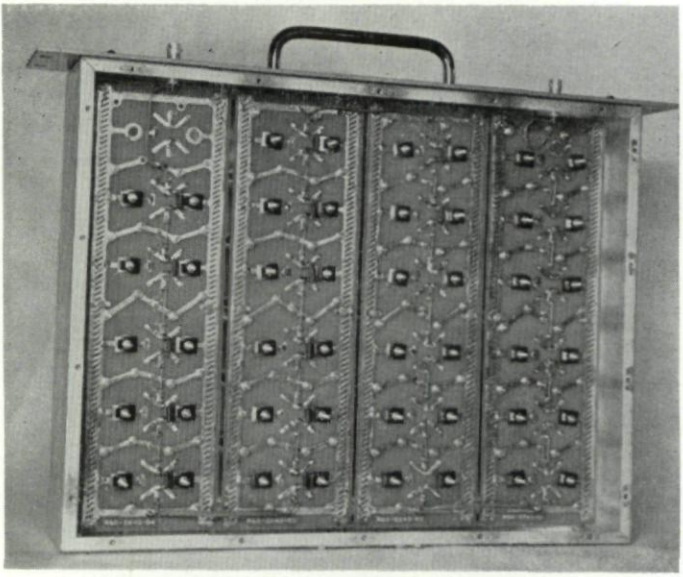


FIG. 20. *A group of sections.*

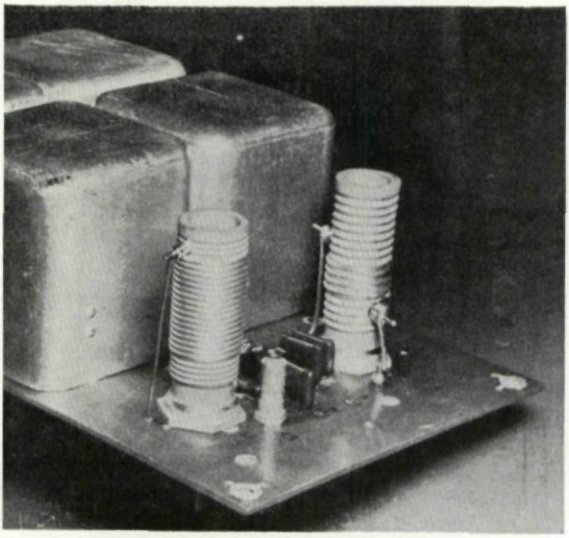


FIG. 21. *A bridged-T circuit assembly.*

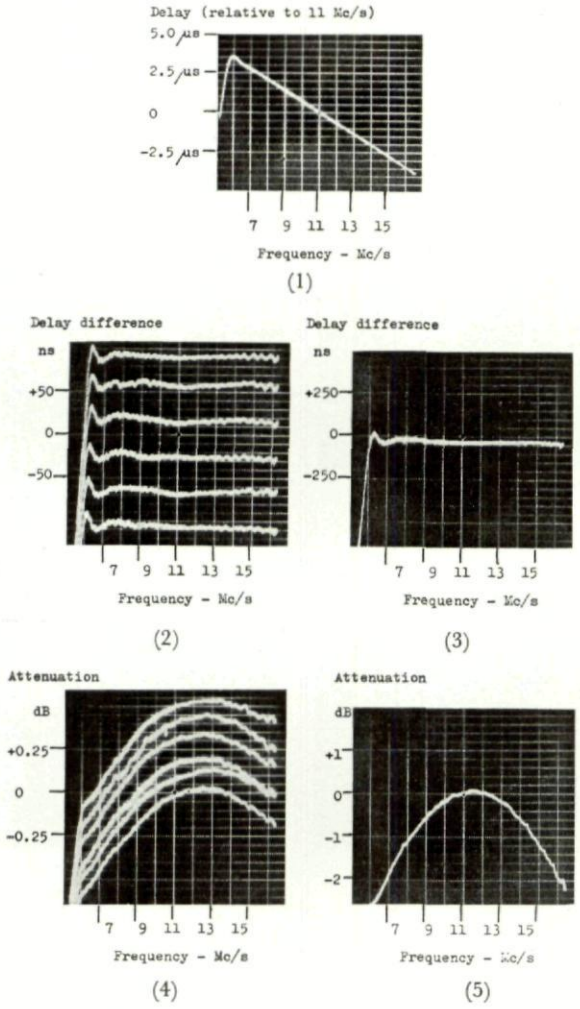
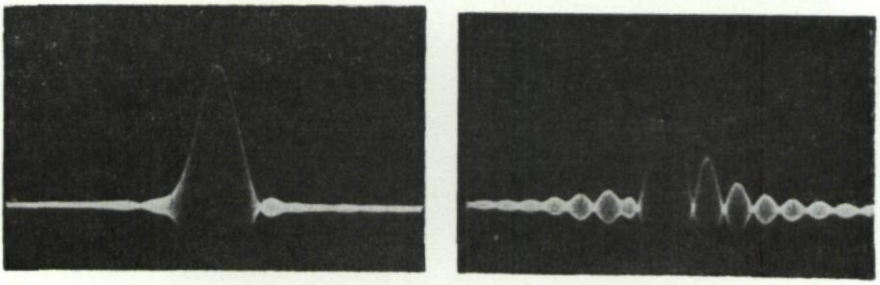


FIG. 22. *The performance of the network:*
(1) *Delay response of network*
(2) *Delay difference response of individual groups*
(3) *Delay difference response of network*
(4) *Attenuation characteristics of individual groups*
(5) *Attenuation characteristic of network*

Due to instrumental limitations the records of Fig. 22 do not show performance above 16.5 mc/s; the delay, delay error and insertion loss of the network are shown at (1), (3), and (5) relative to those at 11 mc/s, whilst delay error and insertion loss of the individual groups are given at (2) and (4) respectively. The delay error characteristics are obtained by generating a saw-tooth shaped voltage waveform corresponding to the ideal or specified

dispersive characteristic and displaying the amplified difference between the ideal and the actual characteristics.

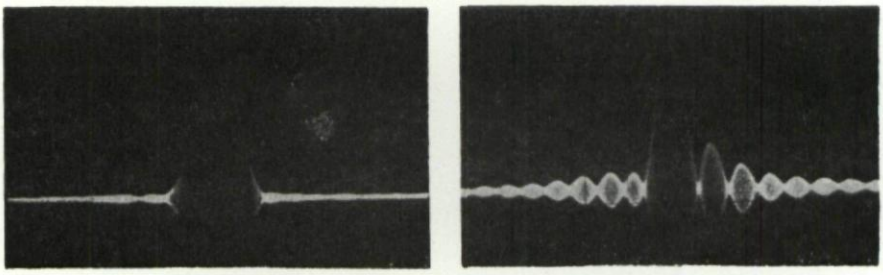
The side-lobe performance of the system is shown in the upper records of Fig. 23 (without compensating section) in which the side lobe level is reduced to a maximum of −26 db as compared with the unweighted response shown on the right. Examination of delay characteristics derived from phase



Weighted Unweighted

DISPERSIVE NETWORK SYSTEM

without compensator



Weighted Unweighted

DISPERSIVE NETWORK SYSTEM

with compensator

FIG. 23. *Side lobe performance with and without compensator*

measurements indicated insufficient delay in the central region of the characteristic so that a compensating section of $Q_e = 1.1$ and centre frequency 15 mc/s was added to the system with the results shown in the lower records where the side lobe level has been decreased to −31 db, compared with the theoretical figure of −34 db when switching transients are excluded. A comparison is shown in Fig. 24 of the measured side lobe levels with and without the compensating section for the complete system with those computed for the measured phase characteristic of the network alone. That this added section has increased the total delay in the region of its centre frequency by about 1% to give a 5 db reduction in side lobe level is an indication of the standard that has to be maintained in the adjustment of individual sections.

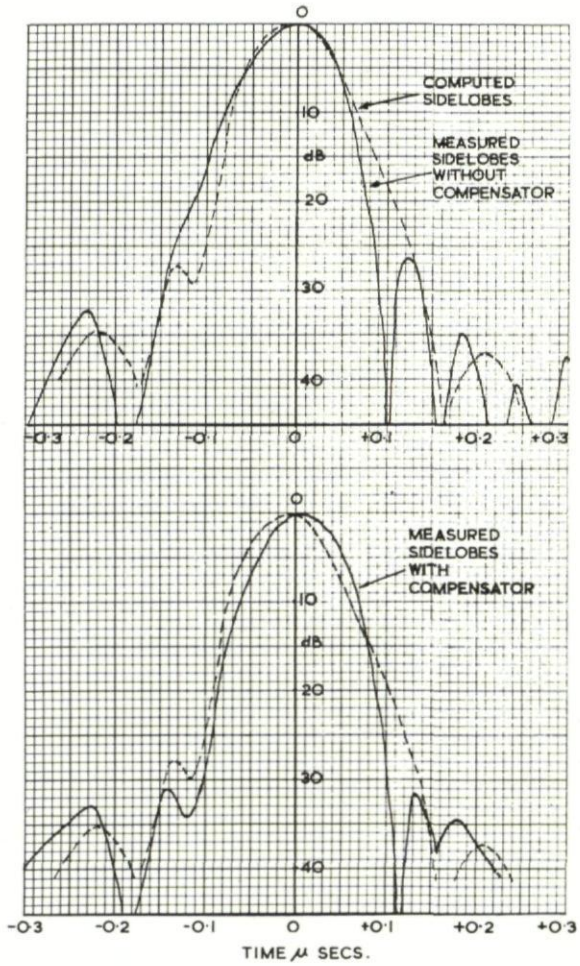


FIG. 24. *Comparison of measured and computed side lobes*

PART 3

DEVICES EMPLOYING ULTRASONIC GRATINGS

W. S. MORTLEY

1. INTRODUCTION

Dispersive properties are also exhibited by mechanical systems and one of the best known devices which have been developed for pulse compression uses the transition between the "thin rod" and "bulk modulus" modes of wave

propagation.[1] The frequency sweep obtainable by this method is limited by the fragility of very thin lines.

Another method[2] makes use of an optical correlation technique in which the signal is put into ultrasonic form in water and used to modulate light in a Schlieren optical system. An optical grating replica of the swept ultrasonic wave is used to intercept the modulated light and a sharp pulse is produced at the moment when wave and mask match one another. This method is limited in frequency sweep by the wavelength in water and the difficulty of producing enough transducer bandwidth. It is also highly inconvenient to have to use a delicately adjusted Schlieren optical system.

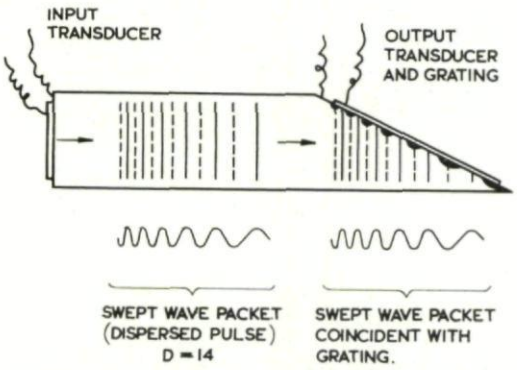


FIG. 25. *U.D.G. device with transmission grating*

In this part of the paper ultrasonic (non-optical) dispersive grating correlators[3] are described with particular reference to a piezo-electric grating (PEG) dispersive device which makes use of ultrasonic propagation in a prismatic block of quartz crystal[4]. It must be appreciated that this kind of device is at a very early experimental stage.

1. *Ultrasonic Dispersive Grating*

The mechanism of ultrasonic grating dispersion is probably understood more easily by first considering the operation upon a short input pulse. Referring to Fig. 25, which shows one arrangement, it will be seen that a pulse travelling from the piezo-electric transducer at the left and impinging upon the grating and the receiving transducer on the sloping face at the right will produce an alternating voltage. This will have a duration equal to the wave propagation velocity multiplied by the length of the transducer measured in the direction of the wave propagation. It will also have a frequency determined by the line spacing of the grating. The spacing may be varied along the grating in any predetermined manner, including that which produces a linearly swept frequency.

It is interesting that this device should produce a swept frequency with a rectangular envelope from a unit input pulse. It does not behave purely as a Fourier transformer, therefore, which would require a $(\sin x)/x$ form of

input pulse for the same output. Even the harmonics could be avoided, in principle, by making the grating with sinusoidal distribution across its width.

Consider now a swept frequency input to the device. As the swept frequency passes the grating it will "beat" with it, producing many cancellations of signals along the grating until the wave group reaches the place where it matches the grating. At this point all parts of the transducer are excited in the same phase and a large, narrow pulse is produced. Clearly the same thing would occur if any other phase code were used instead of the parabolic one, except that the "beats" or "sidelobes" would be different.

Even if we consider only the conventional linearly swept frequency, the

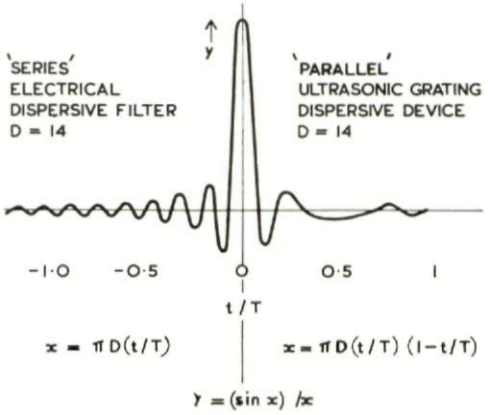


FIG. 26. *Side lobes due to rectangular envelope*

sidelobes of the ultrasonic device differ from those of the electrical dispersive lines discussed in Parts 1 and 2. Fig. 26 shows, on the left-hand side of the vertical axis, the sidelobes from an electrical unit without weighting, fed with a dispersed pulse having a rectangular envelope. These have the usual $(\sin x)/x$ form. On the right-hand side of the axis are shown the sidelobes due to the U.D.G. device, also without weighting, in which x has been multiplied by a factor $|1 - t/T|$ depending upon the overlap of wave-group and grating. It may be seen that the sidelobe spacing is no longer uniform and that the series is limited to a region of twice the dispersed pulse width. The locus of the peak amplitudes of the sidelobes is nearly the same in both cases and the near-in spacings are nearly the same. The differences seem to have no significance.

2. *Reflecting Gratings*

The device described above was chosen for ease of explanation but, although practical for some cases, it has some limitations when the frequency sweep has to be large. There are difficulties in the fabrication of the ultrasonic mask. For this reason one has to consider other methods, of which one is the use of a reflecting grating (Fig. 27) which may be made either by etching lines in the medium or by loading the surface with metal lines.

If we call the angle of incidence of the wave onto the grating θ and the angle of reflection ϕ, the line spacing is related to the frequency (in an isotropic medium) by the equation

$$d = \frac{c}{f(\sin\theta - \sin\phi)}$$

where d is the line spacing
c is the wave velocity
f is the frequency.

The grating is made to conform with the wave group when the latter is incident upon the whole length of the grating.

It may be seen from this equation that if θ is made nearly equal to ϕ, the factor in the brackets is small and therefore f can be made high for a given

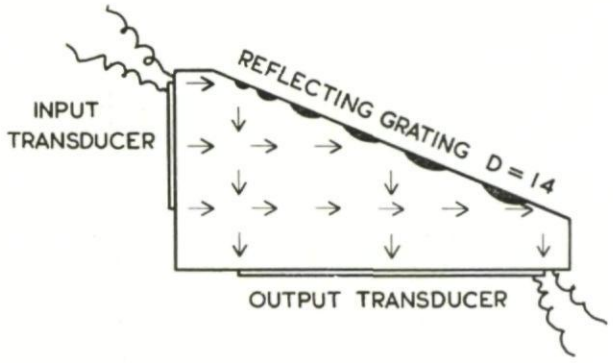


FIG. 27. *Dispersion by a reflecting grating using plane transducers*

practical value of d. This is at the expense of making the device large for a given delay, however, and makes the problem of attaching large transducers more difficult. The transducers can be broken up into mosaics but this also brings difficulties such as the production of ghost signals due to the joins.

Uniform attachment of transducers is very important in this application but it is not obtained by present methods with the regularity which would be required in normal production. This is another objection to the use of mosaic transducers.

3. *Piezo-Electric Gratings (PEG)*

Some of the difficulties of the arrangements which have been described may be overcome by using a piezo-electric medium and making transducers in the form of non-linear combs which also perform the dispersive function. Since there must be both transmitting and receiving transducers, part of the dispersion is done in each and it is simplest to make the parts equal as in Fig. 28. There are interdigitated non-linear comb electrodes on the two sloping faces which launch waves into the medium (crystal quartz) in directions which vary with the frequency and the comb spacing.

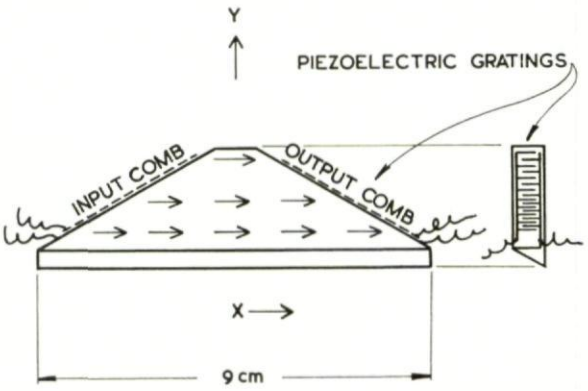


FIG. 28. *First experimental quartz crystal dispersive device*

As the input frequency sweeps downwards, the level in the block at which there is a wave travelling horizontally through the medium rises steadily from the bottom to the top, with steadily decreasing delay. Waves travelling in other directions eventually reach a tilted facet at the bottom which reflects them many times between the two major faces, coated with an absorbing material.

Figure 31 is a photograph showing the pulse compression obtained in the first experimental model. The top trace shows the input dispersed pulse and the lower trace the output compressed pulse, together with some sidelobes and other spurious signals.

Most of these spurious signals are due to a particular choice of wave propagation direction in the quartz. This direction, the *X*-axis, was deliberately chosen for the first experiment in order to help us to understand the results. Its disadvantages are that it produces spurious signals by two

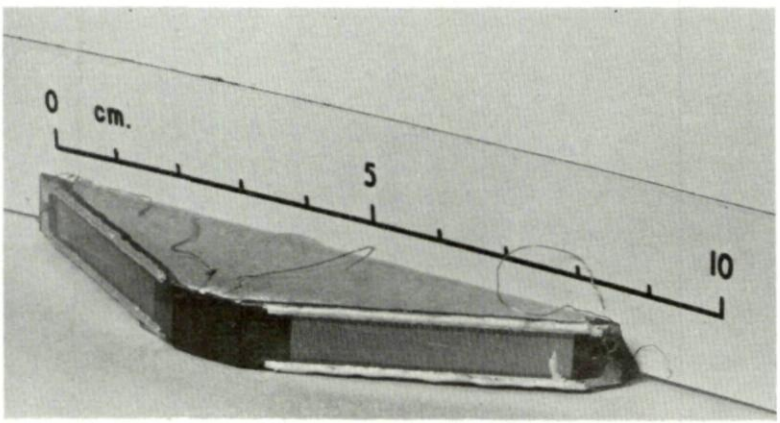


FIG. 29. *Ultrasonic diffraction grating dispersor*

mechanisms, both of which can be avoided by a different choice of direction. One mechanism is through the existence of two waves at velocities which differ by only 12 or 13 per cent. Another is due to a double transformation between wave modes of different velocity. The latter mechanism is a very interesting one but is not relevant to this paper.

Wave propagation in crystalline materials is very complex indeed. Even in isotropic solids there are wave modes with several velocities with transformations between them occurring at interfaces or external surfaces. In crystals there are several elastic constants (six in quartz) resulting in very unfamiliar wave phenomena. For instance the wave energy may be propagated at an angle of up to twenty degrees from the normal to the wave front; or a transverse wave may be transformed to another of the same polarisation but having a different velocity. This transformation occurs in the experimental unit under consideration. Once the wave behaviour is understood, however, it is possible to design a unit which will not suffer in these ways, even with a sweep frequency of half the mean frequency.

This type of crystal device has a great advantage over the previous construction in that there is no problem of sticking transducers on to materials of differing temperature coefficient of expansion. The transducers are merely metal films deposited by evaporation and photo-lithographic processes. Reproduction is simple and the unit is small and robust. The chief disadvantage is that the dispersive delay is limited to about 20 μs by the size of quartz crystals available.

It would be possible to cascade several units with wide-band amplifiers interspersed between them, in order to overcome this difficulty.

Frequency sweep may be increased by reducing the line spacing and by reducing the included angle between the gratings. The experimental unit was designed for a frequency sweeping from 40 mc/s to 25 mc/s, using the longitudinal wave in the X direction. The preferred direction will require a finer grating for the same frequency but, even so, there is just a possibility that a 200 mc/s sweep may be achieved over 5 μs (dispersion factor of 1000) in one unit. The same master negative grating can be used for different sweeps and delays, the dispersion factor (sometimes loosely called the compression ratio) remaining constant. At present it is not known whether the absorption of the wave in crystal quartz will be too great at 500 mc/s or whether the spurious responses will be low enough with small included angles between the gratings.

In the event of such limitations it is possible to obtain a wide frequency sweep by combining dispersive units with coherent frequency changes and added delays but the engineering difficulties are increased especially in regard to phase matching. Phase matching may be achieved by making the electrode gratings separate from the medium and adjusting their positions on the sloping faces with some precision.

4. *Rayleigh Surface Waves*

The comb electrode structure lends itself very well to the generation of surface waves and these have some attractive features. Perhaps the most

attractive is that a simple geometry may be used, both gratings lying in one plane. Another feature is that the propagated wave can spread in only one plane instead of two and waves spuriously radiated into the body of the quartz are very easily separated and absorbed. Yet another feature is that the surface wave propagation velocity is less than that of both the longitudinal and transverse waves. This increases the delay at the expense of frequency sweep.

The chief disadvantage of surface waves is that the colinearity of the gratings limits the frequency well below that which can be obtained with "bulk waves." Another possible disadvantage is that the mass of the gratings will have some effect upon wave velocity and might be difficult to control accurately. On the other hand it might be used as an adjustment.

In spite of the disadvantages it is possible that surface waves would be useful for cheap dispersors, limited to a dispersion factor of one or two-hundred, say, and a frequency sweep of about 30 mc/s.

5. *Signal Noise Ratio and Dynamic Range*

Peak signal to noise power ratio is improved in the ratio of pulse compression, so that 1000 : 1 compression would result in 30 db improvement. A similar improvement will apply to uncompressed signals travelling unwanted paths and if pulse compression is required only to give good time discrimination this is a useful feature. However, if pulse compression is required in order to help to discriminate targets from a background of clutter it is the mean power of the spurious signals which is relevant and the 30 db advantage no longer obtains. For this reason it is essential to reduce this interference as much as possible.

The first step in reducing the level of spurious signals is to deflect them out of the signal path and into ultrasonic absorbers which nearly match the quartz in acoustic impedance (see Fig. 30). In the experimental model there is a mismatch of about two to one, giving a reflection 10 db down. A different material could be used (with greater difficulty) which would improve this to about 17 db. This may not be worth while in view of the multiple reflections which take place.

Some spurious signals reach the receiving transducer directly on sidelobes, however, and cannot be reduced by ultrasonic absorption. They can be reduced by electrically exciting only that part of the array which is sending signals in the right direction. This may be achieved by dividing the electrode combs into groups of teeth which are fed through appropriately tuned circuits. This was not done in the experimental model as may be seen in Fig. 31.

We have already discussed how some of the large spurious signals in Fig. 29 were caused by waves of nearly the same velocity, a longitudinal wave and a transverse wave travelling parallel with one another horizontally across the unit as shown in Fig. 28. This and the late pulse due to double wave mode transformation can be avoided by the methods described above.

Dynamic range is a function of insertion loss when spurious signals have

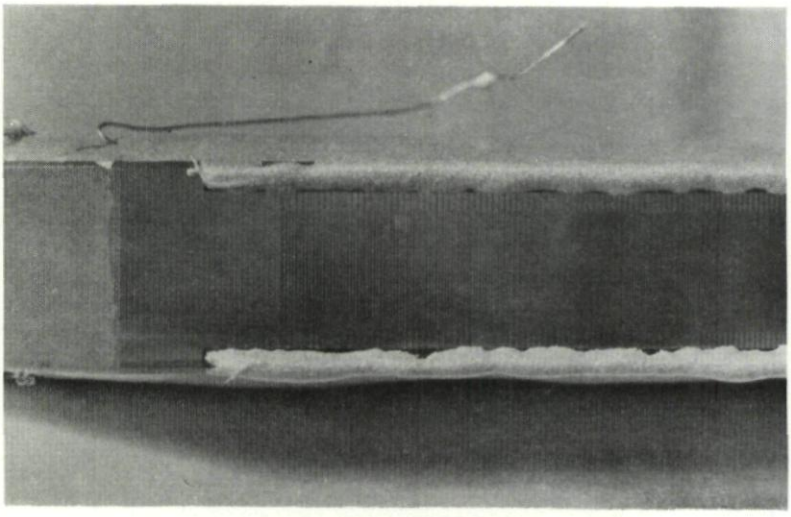


FIG. 30. *Enlarged view of grating*

been suppressed. It is fortunate that the divided electrode scheme for reducing spurious signals will also reduce insertion loss. Improvements of the order of 20 db may be practical.

One other form of interference which can occur in ultrasonic delay devices is a direct electrical coupling between input and output. Such coupling may be kept quite low by the combined use of screening and electrical balance. It is not possible to use very good screening because most of the electrical coupling is through the quartz which forms the ultrasonic signal path. If the ultrasonic absorbers are conducting and grounded, however, the electrical configuration is similar to that of a piston attenuator and stray out-of-balance signals are much reduced.

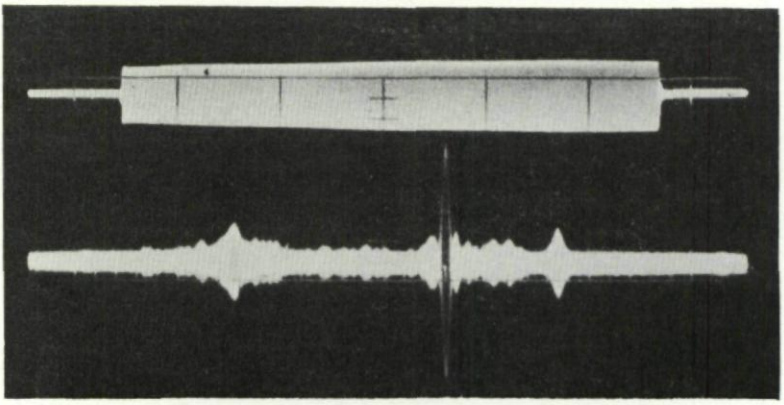


FIG. 31. *Ultrasonically compressed pulse*

6. *Precision*

The precision required of a dispersive unit depends upon its manner of use. For example, in a pulse compression system in which one dispersor is used both for producing the transmitted dispersed pulse (i.e. the swept frequency long pulse) and for compressing the echo signals, first order errors of rate of sweep cancel one another but other errors are added. In some systems it is important to have low sidelobes. In a multi-channel radar system (e.g. one including range trackers), there must be good phase matching between all the dispersive devices used.

For general purposes, therefore, it would seem an advantage to have good precision of rate of sweep and low "periodic errors" (which are responsible for sidelobes). The phase shift at the ends of the sweep relative to the centre frequency is $\pi/4$ times the dispersion factor. If signals here are to contribute to the output positively, then the phase shift must be less than $\pi/2$. The contributions will be 3 db down at $\pi/4$. The precision required, therefore, allowing half the error for the transmission and half for the reception, is $\pi/8$ radians in $\pi D/4$ radians: that is, a precision of mean slope of one part in $2D$. If the dispersion factor D is 1000, for example, the mean slope must be correct to one part in 2000.

In regard to periodic errors, for a small error (or phase or amplitude ripple) the sidelobe produced is half the amplitude of ripple. If one aims at a spurious sidelobe level of — 40 db, therefore, the periodic error must not be greater than 0.02 radian or 2% amplitude. With grating lines π radians apart this means that periodic errors of spacing or width must be less than 1/150 of the spacing or width, respectively. Random errors of spacing and width are of much less importance.

It is seen, therefore, that for high dispersion factors and low sidelobes a high order of precision of manufacture is required. Fortunately this can be met by employing metrological techniques which have been developed for the manufacture of optical diffraction gratings, which involve essentially the same problems on a very similar scale[5]. Very precise angles between the gratings are required, and this can be assured by the use of well-established optical manufacturing techniques.

Another aspect of precision is stability. With metal gratings deposited directly upon the surface of a block of quartz there is no danger from vibration or shock but there is, of course, a variation of both line spacing and wave propagation velocity with temperature. It is expected that the temperature coefficient of dispersion will not be more than a few parts in a million per degree centigrade (at normal temperatures) in the preferred orientation. Stability and low attenuation at high frequencies are the reasons for choosing quartz as the medium, rather than a ceramic.

CONCLUSION

Ultrasonic dispersive grating devices, although not having the dynamic range of electrical filters with dispersive delay, promise to be interesting from the points of view of small size, robustness, repeatability and precision,

especially for large dispersion factors. The PEG device also has the merit of a simple and reliable form of construction.

REFERENCES

1. MAY, J. E., Wire type dispersive ultrasonic delay lines; *Trans. I.R.E.*, **PGUE-7,** 53–58, 1960.
2. SLOBODIN, L., Optical correlation techniques, *Proc. I.E.E.E.*, **51,** 12, 1782, 1963.
3. Marconi Company. Patent applied for 1962.
4. MORTLEY, W. S., Pulse compression by dispersive gratings on crystal quartz. *Marconi Co. Res. Rep.* **RD/P.1694** 1963.
5. Messrs Hilger and Watts, Ltd., London.

DISCUSSION

————: Have you considered the utilisation of the dispersive effect of a truncated waveguide?

P. S. BRANDON: Yes, but the waveguide length is prohibitive for the ranges which we require.

H. MIEDEMA: Can you quote a loss figure for the ultrasonic delay line?

P. S. BRANDON: The loss was about 60–65 db, but there is no doubt that the performance of this early model can be much improved.

CHAPTER 27

RADAR DATA-HANDLING AND DISPLAY SYSTEMS FOR USE WITH PULSE RADARS

A. P. YOUNG
The Marconi Co. Ltd., Chelmsford, England

The purpose of a radar system is to gather data from radar and other sources and to use them to provide effective control for civil traffic or for military aircraft and weapons, or for the two in combination. Although the ultimate objectives in civil and military environments are dissimilar the methods of instrumentation would typically have much in common.

Figure 1 shows in outline the flow of data in a military system. Raw radar signals are fed to displays at appropriate operator positions and are used by tracking operators to initiate the tracking of aircraft. Tracking is carried out using markers pointed by electronic means on the face of the P.P.I., the operator positioning the marker by means of a tracker ball control normally situated at the right-hand end of the desk. The resulting plan-position data may be used to control height-finding and other equipment, while push-button controls and typewriter keyboards allow additional information to be entered into the computer, thereby building up a detailed account in it of tracks, of weapon states and assignments and of desired procedures. This

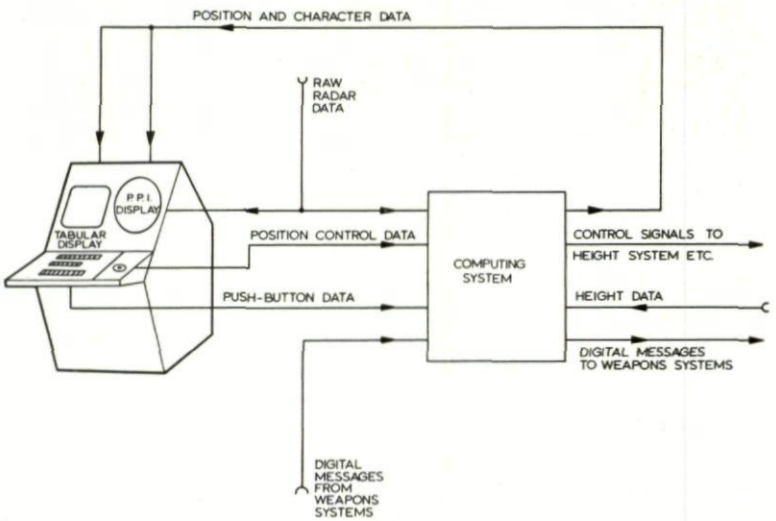


FIG. 1. *The flow of data in a military system.*

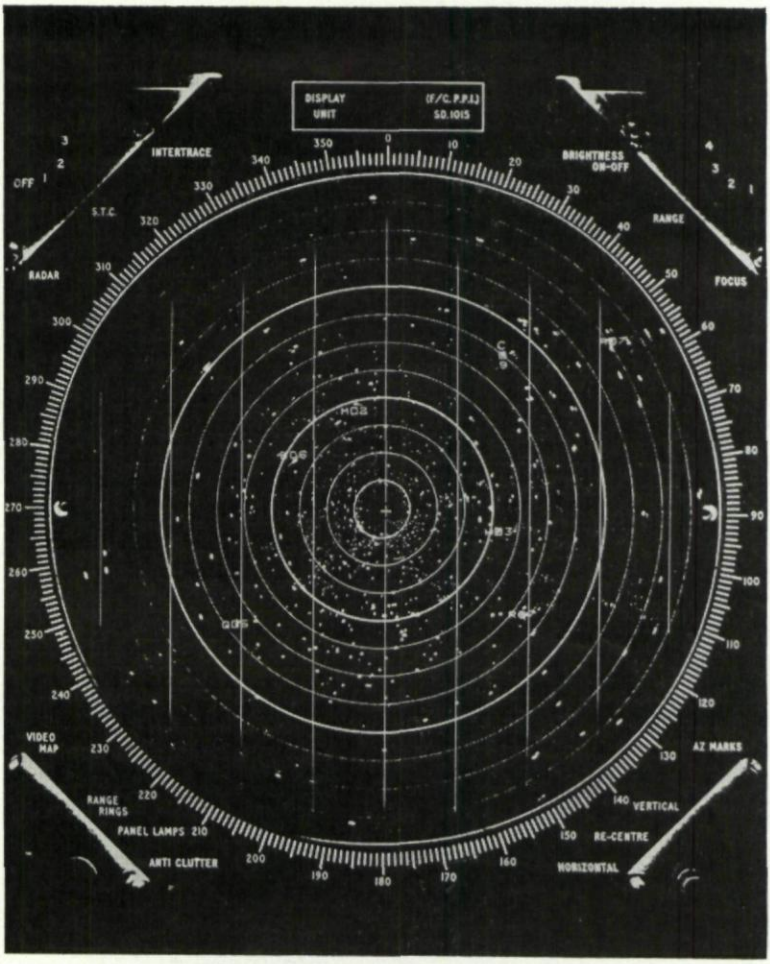


FIG. 2. *A marked raw radar display.*

information may be supplemented by digital messages from other radar centres and from weapons systems.

Track data and control information thus derived are fed to P.P.I. and tabular displays at the various operator positions, and may also be fed by digital data link to other radar centres, to aircraft and to weapons systems. Figure 2 shows a Marked Raw Radar Display, horizontal markers representing aircraft tracks while the vertical marker might be used for pointing (inter-console marking) or by a tracking operator for tracking. Figure 3 shows a "synthetic" P.P.I. picture, in this case a possible civil air traffic configuration. Figure 4 shows a Cathode Ray Tube Tabular Display of general Aircraft Track Data in a military system, five tracks having been selected by the operator. The bottom line has been used in this case to

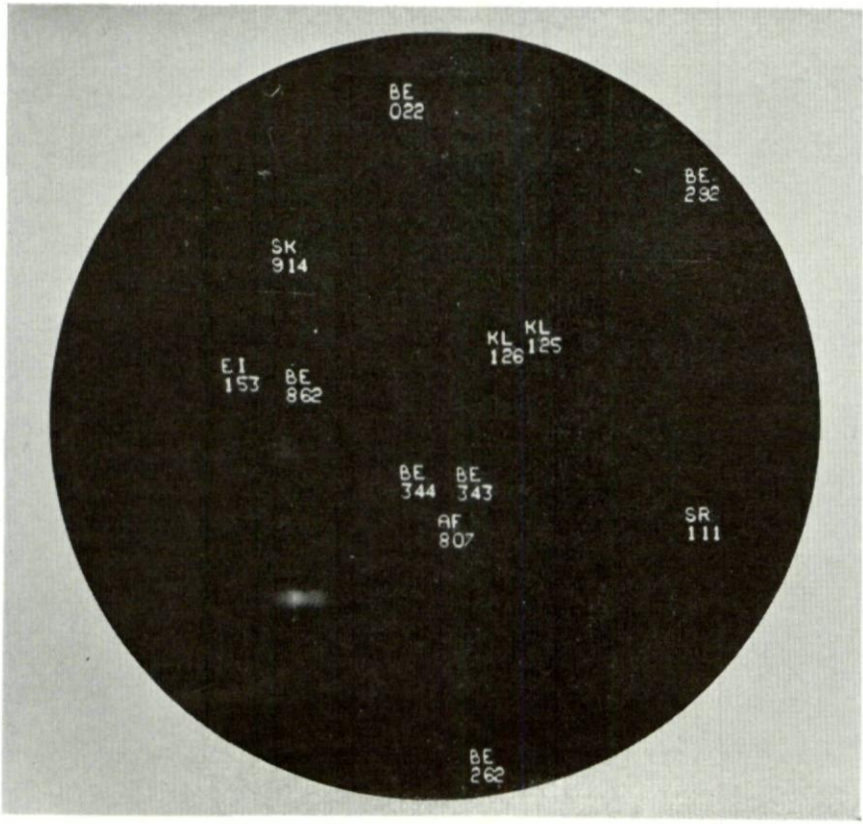


FIG. 3. *A "synthetic" PPI picture.*

TRACK NUMBER	CALL SIGN	IDENT	SPEED	HEADING	HEIGHT	WEAPONS	STRENGTH
A4	C135	F	650	123	275		
B6		X	530	256	330	F	
R7		H	560	274	300	S	4
R9	SW3	T	286	198	210		
S3			673	147	360		

S3-ID-H-STR-1

FIG. 4. *A cathode ray tube tabular display.*

provide tell-back to the operator in respect of push-button controls, and shows that the computer has received the message "Track number S3-Identity Hostile—Strength 1." This instruction may now be "registered" by the operator, in which case the identity and strength would be recorded in the computer's track store and would appear in the S3 track data, the tell-back line being extinguished. In case of error the operator would cancel the tell-back line.

Figure 5 shows a comprehensive display suite for the use of a weapons

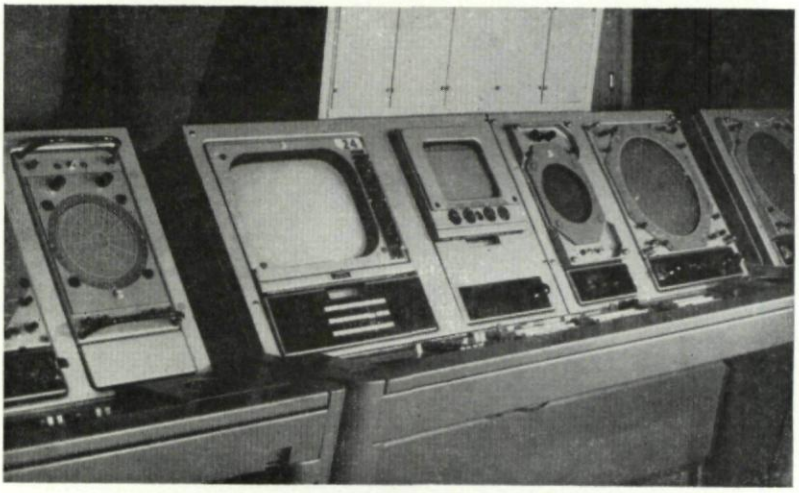


FIG. 5. *A comprehensive display suite.*

control officer and his assistant, and contains, from left to right, a closed-loop television system for showing weather maps and semi-static information, a tabular display for weapons supervision and control data, a monitor display and a marked raw radar display.

Track initiation and marker positioning are carried out using a Tracker Ball control shown mounted at the right of the desk of Fig. 6. The mechanism is shown in Fig. 7. Rotation of the ball is resolved in two directions at right angles by two shafts in contact with it. Each shaft carries a castellated disc, one of which can be seen in the illustration, and these are used to sense the magnitude and direction of movements, two photocells and two lamps being used. The resulting signals are fed to the computing system where they are used to modify the marker position. The control has the advantage that there are no mechanical "end-stops" so that it can be used in an incremental manner without need for mechanical recentering.

When used for track initiation and maintenance, the tracker positions the marker over successive echos and presses a button, causing the position of the aircraft to be recorded, the readings being used in the computer to form and to up-date the position, speed and heading of the aircraft. The computer will normally be programmed to minimise the effects of errors in these

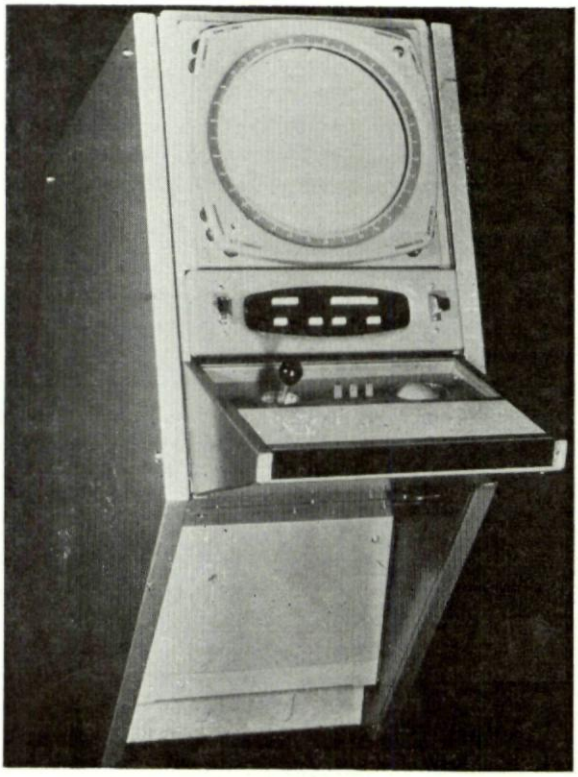


FIG. 6. *A tracker ball control.*

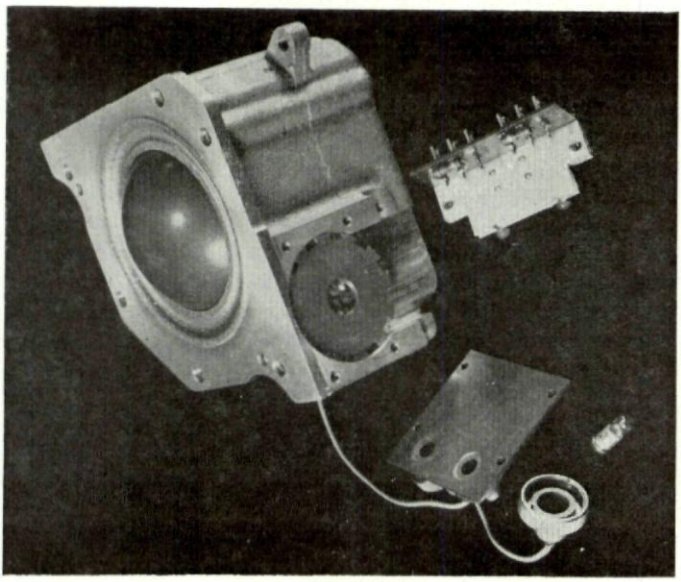


FIG. 7. *The tracker ball control mechanism.*

readings, it will generally apply different smoothing factors "along" and "across" track, and reduce the degree of smoothing used when turns are encountered. In the case of co-operating aircraft the tracking accuracy can be further improved by using knowledge of expected track behaviour in the tracking programme—a particularly valuable resource in the case of fighter aircraft.

When track position, speed and heading have been established the position of the track marker is continuously up-dated by the computer, so that intervention of the tracking operator is necessary only when an aircraft changes course or when an appreciable error is observed. However, the task of the tracker is much simplified, and the tracking accuracy is improved, by the use of a system known as "hopping strobe," whereby the rate-aiding of the tracking marker takes place in a series of hops, each hop occurring at the instant when the radar scans the followed aircraft. This procedure removes errors which would otherwise appear due to variations in the time at which the correction is recorded.

Thus by using push-button desk controls in association with the rolling ball position control a tracker is able to initiate and to maintain a number of aircraft tracks. A tracker may also use these controls to cause an established track to be maintained automatically by the computing system, echo positions being measured by electronic means and used in the computing system to up-date the track, thus increasing both the tracking accuracy and the tracking capacity of the individual operator. The "double-threshold" method of echo-detection has been successfully employed for this purpose, and films show first the efficiency of the detection process, and second its use in obtaining automatic follow. The films were taken using one frame for each radar head rotation, and are projected at a rapid rate to give a speeded-up presentation. In the first film bright dots appearing near the echos and following them in azimuth, show successful recognition of the echo. In the second the use of these data to maintain tracking continuity is illustrated.

Turning now to the general design problem, it will be clear that a computing system must be designed to accept signals from all data sources, and to deliver signals to all output devices simultaneously. The methods used to achieve this vary widely according to the capability of the digital computer used. With slower computers the system designer is forced to use a diversity of specialised and relatively complex mechanisms in conjunction with the general-purpose machine, to obtain the necessary speed of input/output and of processing. As the capability of the computer increases, the amount of specialised equipment is reduced, and more and more of the functions are absorbed into the computer. This has many advantages of which the principal ones are:

(1) The overall system cost is reduced.

(2) The system becomes more flexible.

(3) Maintenance is simplified and the equipment content is reduced.

For example, Fig. 8 shows a block schematic of the Tabular Display

System. Here the character number is derived from the store under the control of the address counter, and is "decoded" to produce a series of instructions leading to the drawing of the symbol in the manner shown in the bottom left-hand corner of the picture. With a well organised high-speed computer it is possible to use the computer to replace the store and address counter. Again where outputs to high-speed digital data links are involved the use of low speed computers will normally demand the use of output buffer registers in conjunction with parallel-to-serial conversion equipment, to allow the computer to put out the message a number of bits at a time.

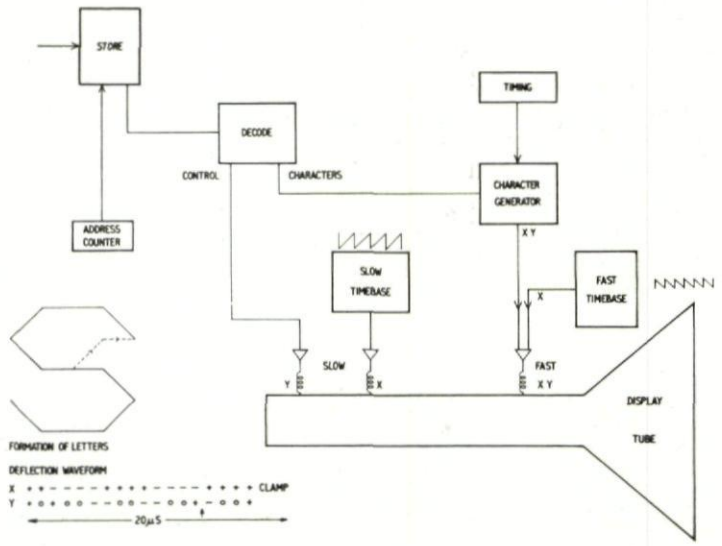


FIG. 8. *Tabular display system.*

Improved computer capability can reduce this equipment by using the computer to put out the message one bit at a time.

To give some indication of the progress being made in the design of computers for this class of work, Figs. 9 and 10 are included. The first shows the TAC computer, designed in 1959. It is about 7′ high by 10′ wide by 1.5′ deep, has a store cycle of 10 μs and executes most short orders such as "add single-length" in 22 μs. The new *Myriad* computer, of which an assessment model is shown in Fig. 10, is about 41″ high overall, by 4′–8″ wide by 2′9″ deep. It uses a 1.2 μs store cycle (one core per bit) and the corresponding short order time is 2.5 μs. Experience to date suggests that the operational advantages of the new techniques employed in this machine are accompanied by a significant improvement in reliability. While this is, of course, an extremely important aspect of equipment performance it is certain that in many applications stand-by equipment, in one form or another, will continue to be demanded by the user because continuity of service is essential. It is, therefore, particularly important to minimise the cost and

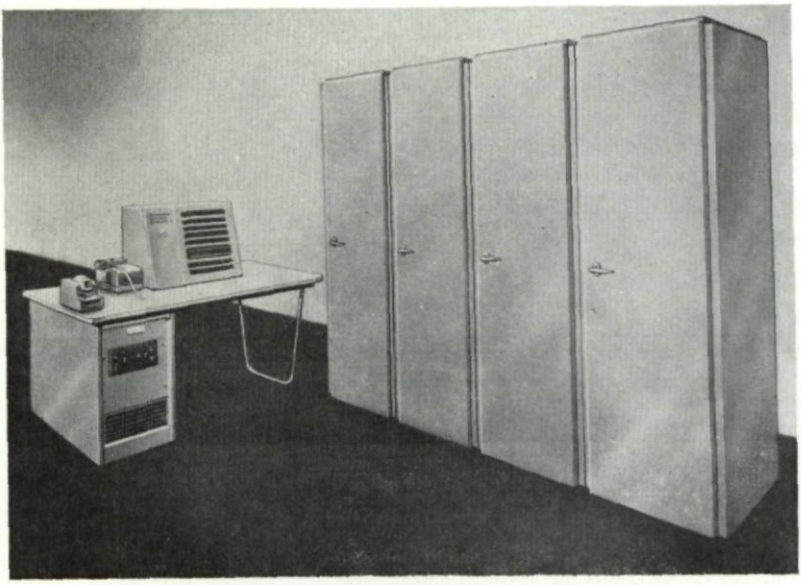


FIG. 9. *The T.A.C. computer.*

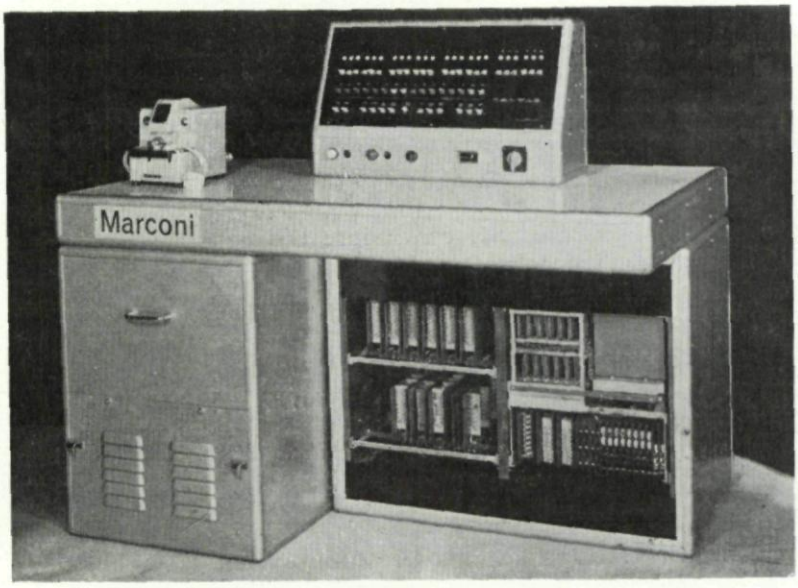


FIG. 10. *The Myriad computer.*

complexity of the computer, obtaining high operating speeds by skilful circuit design and by relatively simple organisation rather than by a proliferation of equipment. This approach, coupled with the use of a minimal amount of peripheral equipment, also improves the ability to detect failures automatically when they occur, and to initiate change-over procedures.

DISCUSSION

P. V. ASHWORTH: You mentioned that the main advantages of the ball controller was that it had no end stops. If you had used the joy stick as a rate controller you would also have no end stops. Have you explored this avenue? In other words is there an advantage to be gained as a position controller rather than a rate controller?

A. P. YOUNG: I think that we have actually made both forms of machine. When you are tracking you want to apply directions to a number of aircraft tracks in succession, you have a sequence button so that you can transfer the control of the tracker ball from one aircraft to the next and so on in some sort of sequential way. Now in this sense ideally when you press the sequence button the marker will come up with the predicted position of the next aircraft in the sequence and you wish really only to make position corrections, you don't really want to use a rate control because the error is going to be small. Now, the difficulty arises with the joy stick that if you have to make 20 successive corrections all northwards for example, the joy stick would just go off the end of the travel.

E. KEONSIAN: You have referred to the fact that fast computers are much cheaper. I presume that this is due to the fact that the peripheral equipment which goes with that computer is simplier, therefore you presume that it should be cheaper. How about the basic computer itself. We have experienced that usually very fast computers are also very very expensive.

A. P. YOUNG: It depends upon how you get the speed. If you get the speed by an elaboration of equipment rather than by being very clever, then I agree with you the computer does cost more, but if you get your speed by clever circuit design by taking advantage of all the possibilities that the market offers, I think in that case you can produce computers which are in the region 10 to 20% more expensive and perhaps five times faster.

CHAPTER 28

DIGITALISATION OF RADAR SIGNALS AND THEIR EVALUATION BY A COMPUTER FOR AUTOMATIC TRACKING OF TARGETS

K. v. Schlachta, H. Springer, W. Storz, W. D. Wirth
Institut für Funk und Mathematik, Germany

This paper contains a description of a technical system intended for real-time evaluation of surveillance radar-signals for observation of air-space.

This development was initiated as part of an attempt to automize civil air-traffic-control. As the digital radar system has only the task of observation, the feed-back towards the observed object, that is the aircraft, must be done by an air-controller. It is planned to extend the system to a general automatic collision warning.

As the echoes of the surveillance radar cannot enter the computer directly, they must be processed. The technical device for this, the so called digital detector, was constructed by Dr. W. Storz and Dr. W. D. Wirth under the supervision of Prof. Dr. F. W. Gundlach and Dr. H. Jungfer in the Heinrich-Hertz Institut (Berlin).

The "detector-data" edited by the digital detector are processed by a computer which must perform several tasks as a controller: storing, associating, and identifying signals. This radar-data processing philosophy was developed in the Hahn-Meitner-Institut (Berlin) under the supervision of Prof. Dr. W. Haack by Dr. E. Jessen and Dr. H. Springer and their computer-team.

In our description of the digital radar system we follow the functional connection of the different parts of the system, beginning with the description of the tasks and the logical operations of the digital detector. Then we describe the computer programme for radar-data processing. Emphasis is given to the tactics of automatic target tracking and those of continuous acquisition of new targets.

Surveillance-radar-systems radiate via the rotating antenna short pulses with a constant pulse recurrence frequency. Scanning a target, a group of echo signals is received, which because of their constant delay against the radiated pulse can be attached to a distance ring. The width of such a ring is chosen $c\tau/2$ corresponding to the width τ of the radiated pulse. For $\tau = 2\ \mu s$ this results in a width of the distance rings of 300 m.

The area observed by the radar is divided into distance rings. Each ring is searched for targets by a detector device following the rotation of the antenna. For the present it is sufficient to consider the processing on one distance ring. The echo signals are added to the receiver noise, the resultant

signal therefore being random. The probability density distribution is a function of the amplitude of the echo signal. With a special decision rule it must be decided by inspection of the signal if there is a target or only noise. By statistical decision theory an optimal processing system can be evaluated. Optimal means, that for a given probability of false alarm by noise the probability of detecting targets is a maximum in comparison with any other processing system. For optimal processing of radar-video-signals, the echo signals and noise of a target from one scanning must be weighted by a weighting function derived from the antenna radiation pattern, and added up. If the sum is higher than a given threshold, a target is assumed to be present. Such an optimal system can only be realized by storing many analog signals. But it gives also a measure of efficiency to a simplified system.

Quantizing of the video-signals in 1 bit (0 or 1) we get the advantages of digital processing and storing. If the video-signal exceeds a quantizing threshold, a 1-pulse is produced, otherwise a 0-pulse. The 0- and 1-pulses are arranged into a 2 μs-clock rate synchronized on the radiated pulse for the division of the range in distance rings. The reduction of range of the whole system against an optimal system by quantizing the video-signals is estimated by a computed comparison to 4–7%.

For optimal processing of the quantized video-signals, the group of echo signals with noise must also be weighted by a weighting function derived from the antenna radiation pattern and then added up. The sum has to be compared with a threshold, too. By a computed comparison it may be shown, that without reduction of range the weighting function can be replaced by a rectangular one, with the width corresponding to the half power beam width of the antenna.

This results in the well-known moving average or sliding window detector, which after this is as good as an optimal digital detector.

Our experimental system contains the further simplified pulse-pattern-detector. This detector proves to have a reduced range of 5–10% compared with the optimal digital detector. In this pulse-pattern-detector a binary counter with 5 stages is controlled by the actual digital videosignal and by the digital videosignal delayed one radar period. If one or other of the two digital videosignals is 1, the counter counts upwards. If both of them are 0, the counter is reset, if it has not yet exceeded a detection threshold j_1. Therefore the counter continues to count if in a 0-1-pattern of the digital video-signals there are single zeros.

The antenna angle with respect to north is taken as the target azimuth. This corresponds to the centre of the echo signal group. For the pulse-pattern-detector as target-centre, the average between starting the detector counter at the target-beginning and stopping the counter at the target-end must be taken. The target-centre angle can be computed from azimuth at target-end and from the position reached by the detector counter at target-end as a measure for the target width. In our system the target-centre is estimated by doubling the counting rate after the target-end. When the detector counter reaches 31 (11111) a target pulse is produced with a constant delay of 16 radar-periods to the target centre. If the detector

counter before target-end exceeds a threshold j_2 higher than the detection threshold j_1, a strong target is assumed to be present and a special mark signal is produced with the target pulse. This signal facilitates the primary registration to the following tracking computer.

For the delivery of a 13 bit azimuth word the position of a binary counter is gated by a target pulse parallel to the output. This counter is reset by a north pulse and is set by azimuth pulses from a coding disk, geared to the antenna.

The processing rate for a distance ring equals the pulse recurrence frequency but it is delayed corresponding to the distance. For the estimation of the distance the target pulse produced by the detector for a distance ring gates the position of a distance-counter, which is reset by the p.r.f.-trigger and then is set at a 2 μs-rate. The whole target word comprises 10 bits for the distance, 13 bits for the azimuth and 1 bit for marking strong targets, all together there are 24 bits.

For processing all distance rings of the area observed by the radar we applied a multiplex system consisting mainly of one detector device and a core memory (Fig. 1). In the core memory a word for each distance ring is provided. For the actual distance ring the detector receives the provisional result from the period before, processes it together with the actual digital videosignal and delivers the new provisional result to the memory.

The succession of distance rings is processed at a 2 μs-rate. The available core memory, however, has an 8 μs cycle time. To adapt the two rates an intermediate buffer register is necessary. This register accepts the provisional results or information of 4 distance rings in common from the core memory and delivers them separately to the detector.

The detector consists of 4 register columns with intermediate And-Or gates. The information of a distance ring is shifted parallel in a 2 μs rate through these register columns. During this shifting the decisions are performed by the logic network. The provisional result of the last period is shifted together with the actual digital videosignal into the first register column. During the shifting process the new provisional result is produced. By this shifting process the detector can receive new informations at a 2 μs-rate though the processing itself is performed at a 2 μs-rate, too. In the intermediate buffer register the detector outputs are gathered for 4 distance rings and then are delivered together to the core memory. In this way the circulation of the information is completed.

The logic decisions are arranged into the shifting process in the following way. Shifting from RI to RII it is checked by the actual and delayed digital video signal, whether the detector counter is to set one step upwards. Also checked is, whether the counter reached 31 (11111) for producing a target pulse, whether at target-end the counter exceeds the threshold j_1 and therefore has to count at a doubled frequency, and whether the detector counter is to be reset. Shifting from RII to RIII a one is added to the counter position, if it must be counted. If the counter is to be reset, 00000 is put into RIII for the counter position.

To control all logic functions of the detector, a test signal is produced on

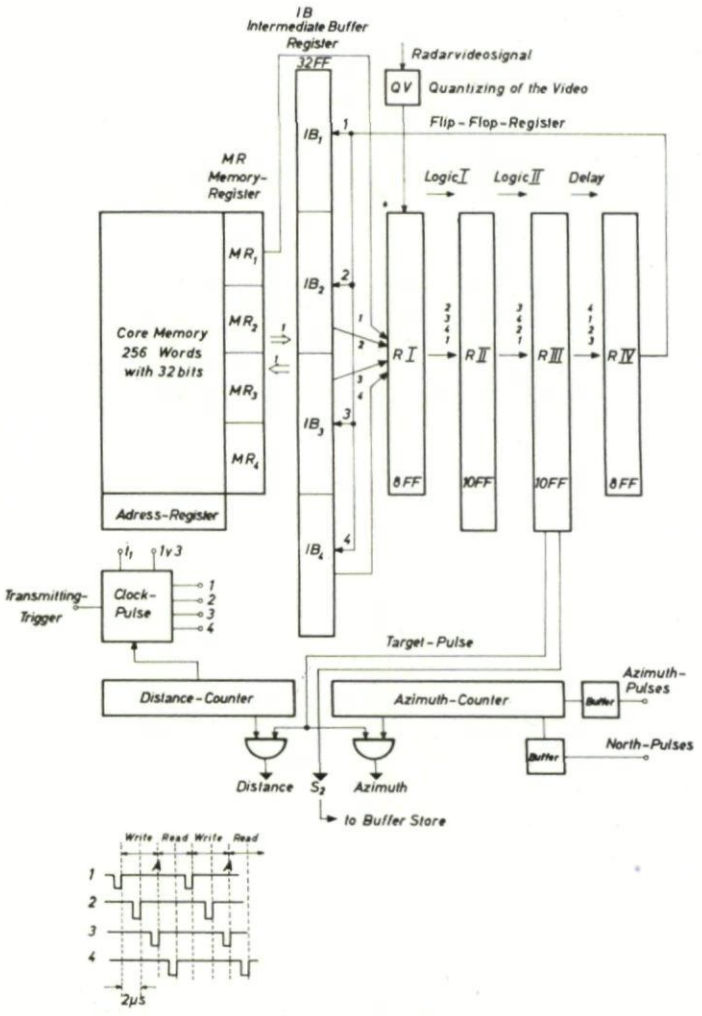


FIG. 1. *Digital detector multiplex system.*

the outermost distance ring which is blocked for the radar video signal. The test signal is sufficient to the threshold condition j_1. If the target pulse is not produced or is produced at a wrong time, an alarm is given.

The target pulses, delivered by the multiplex detector for all distance rings, induce the delivery of the target words. These target pulses of the detector occur irregularly. Their minimum distance is 4 μs. If the target words should be given on a channel of small bandwidth, for example a high speed tape puncher or a telephone line, then the flow of information must be smoothed by a buffer store. By this means the bandwidth becomes considerably smaller. We assume for example a number of 600 true and

false target pulses per antenna revolution and 12 antenna revolutions per minute. Then the average time interval between two target pulses is 10 ms. This average time interval is greater than the minimum distance of 4 μs by a factor of 2500. The bandwidth of the output channel therefore can be reduced by a buffer store by the factor 2500. With a target word of 24 bits the information rate after the buffer store is 2.4 bits per ms in this example. The buffer store has the following construction (Fig. 2): the storage is made

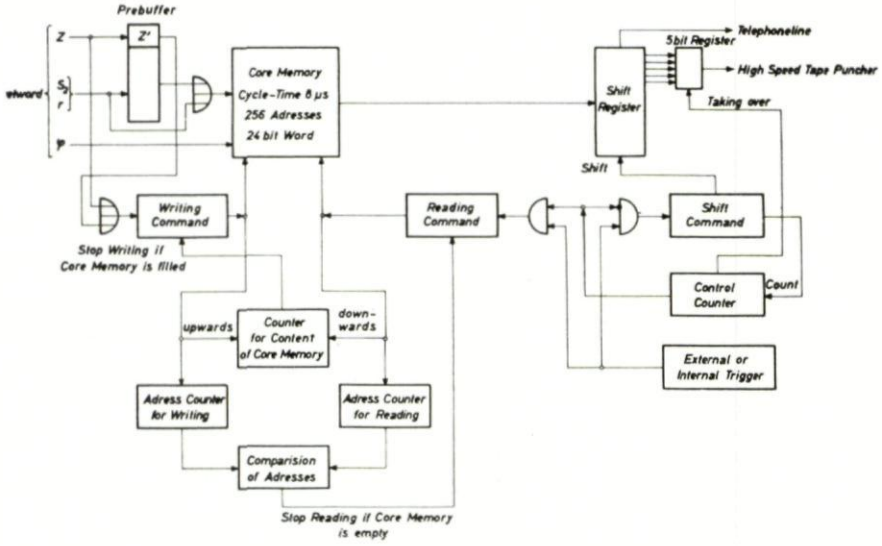


FIG. 2. *Block diagram of the buffer store.*

with a core memory with a cycle time of 8 μs, 256 addresses and a word length of 24 bits. Since the target pulses are synchronized with a 2 μs clock rate, the target words must be arranged in the core memory cycle of 8 μs by a prebuffer consisting of flip-flops.

The target word is delivered out of the buffer store by a shift register either in series or to a high speed tape punch in groups of five bits. The core memory has two address counters, one for writing and one for reading. The corresponding address is increased by one with each writing or reading. The command for reading is stopped, if the two addresses are equal, preventing the address for reading to exceed the address for writing. A further counter counts the targets stored in the core memory. This counter takes care that the address for writing does not outrun the address for reading, preventing targets still stored being written over. A target word is written in the core memory, if there is a target pulse or a target stored in the prebuffer in the moment of the writing-cycle. A target word is given to the output by an internal or an external trigger. A reading command is produced by the trigger, if according to the control counter the target word is shifted out of the shift register and if the core memory is not empty. A shift command is

produced, if according to the control counter there is still a target word in the shift register, which is not completely shifted out of the register. The buffer store is checked permanently. Check targets are produced at the input of the buffer store and it is checked whether these check targets appear at the output.

The number of target pulses per antenna rotation varies by a long-time fluctuation of the radar video and of the threshold of the digitalisation. A long-time control equalizes this fluctuation and keeps constant the number of target pulses per antenna rotation by varying the threshold of digitalisation, which means by varying the sensitivity of the detector.

The normal values of the radar video, the threshold of digitalisation and the threshold of the detector may be so, that the device at the output of the buffer store, for example a high speed tape punch, a telephone line or a digital computer, operates with maximum speed of receiving except for a little in reserve.

An increase of the radar video for example causes an increase of the number of target pulses per antenna rotation. Then the device at the output can no longer receive the average information rate. So the buffer becomes completely filled and target pulses are lost. If the radar video decreases, the buffer store is frequently empty and the device at the output cannot operate with maximum speed. Therefore the long-time control takes as controlled condition the average filling of the buffer store. The content of the counter, which counts the number of stored target words in the core memory, is averaged. The average is compared with a nominal value and the threshold of digitalisation is controlled by the difference of these values.

If a digital computer is at the output of the buffer store, it is possible that the computer can receive more target words than it can process in real time. In this case the number of target pulses per antenna rotation is not controlled by the average content of the buffer store. The target pulses per antenna rotation are counted by a counter. The content of the counter after each antenna rotation then controls the threshold of the digitalisation.

A fast threshold control, in function with the distance, raises the quantizing threshold for the video signal in clutter areas. The video signal quantized by an uncontrolled threshold is counted in distance segments of 8 distance rings. The counting result is delayed by the core memory of the multiplex detector for one radar period and serves after a digital-analog conversion as the quantizing threshold in the corresponding distance segment This fast controlled threshold is multiplied by the control voltage of the long-time control for the stability of the system.

Simultaneously with the registration on punched tape, the target pulse was added to the video signal for the PPI. Adjacent to each target blip the target pulse is visible as a short radial line. The series of PPI pictures was photographically recorded. For each detector adjustment a series of about 200 pictures was produced. With these pictures we later on compared the target detection by the detector with the blips. The series of pictures was projected on a sheet of paper and the targets tracked by painting. It was noted whether as a target was detected the blip was visible on the PPI and

vice versa. In Fig. 3 you see the results. Of most interest is the probability of target detection in comparison to the probability of seeing a blip. The results of our experiments show that our relative simple detector is at least equivalent to a human observer.

For the evaluation of a detector-logic, which is suitable for the conventional MTI-system, we recorded the echo-signals of single aircraft by means of an oscilloscope and a camera. The video signals of moving targets show typical reductions of pulse aplitude during scanning by the antenna. These periodical reductions of pulse amplitude find an interpretation through

Target-End Condition Detection Threshold	False Alarm Probability	Detection Probability provided Visible Blip on the PPI	Detection Probability provided Unvisible Blib on the PPI	Detection Probability provided a Target Is Present	
				Detector	PPI
$V_1 = 0$ $j_1 = 5$	10^{-5}	0,924	0,224	0,63	0,575
$V_1 = 0$ $j_1 = 12$	$5 \cdot 10^{-6}$	0,974	0,174	0,80	0,78
$V_1 \vee V_2 = 0$ $j_1 = 5$	$6 \cdot 10^{-6}$	0,989	0,424	0,81	0,68
$V_1 \vee V_2 = 0$ $j_1 = 12$	$3 \cdot 10^{-6}$	0,828	0,031	0,64	0,77

FIG. 3

the theory of the MTI-receiver in an ASR-system. The output of the MTI-phase detector is a function of the phase difference between the echo signal and the transmitted signal. The difference between two successive pulses gives the video signal. The output voltage of the phase detector depends on the cosine of the phase angle. For this reason the difference between two successive pulses can be zero without the phase difference being a multiple of 2π.

There exists a relationship between the period and reduction of amplitude and the velocity of the target. Also the phase at the beginning of the pulse group from one target effects the amplitude reduction. If the variation of the phase between two pulses is small, it is possible that the amplitude goes to zero during as many as ten of the pulses. With a low probability the group of received pulses is divided into two or more parts.

By recording the video signal and simultaneously calculating the radial velocity from the coordinates of the digital detector there was found a connection between the theoretical and the recorded periodic reduction of amplitude. We see from this, that the uniform leap of the phase in the signal reflected by an aircraft is not super-imposed with a significant fluctuation. It is the problem of the digital detector to continue counting if there are intervals between the echo signals. Otherwise the target would not be detected or there would be a splitting into more than one target.

The data edited by the detector are the polar-coordinates of the echoes. The number of data per revolution of the antenna is regulated by the threshold regulation of the digital-detector. In spite of the great amount of preparatory work done by the detector not every signal is the coordinate of a target nor is every target causing a signal. Noise and clutter are causing loss of true and gain of false signals. The task of the computer is to find the tracks of the targets out of these "anonymous" coordinates. The basis of the logic of the surveillance-system is the fact that the coordinates of targets taken during several aerial rotations are correlated thus being distingushed from the noise-signals which are not. The storage enables the computer to analyse the coordinates of several revolutions while the digital detector just analyses the situation during one aerial rotation.

The surveillance of the airspace requires the acquisition, identification and tracking of all targets by the computer in a similar way to that in which it is done by the controller. We shall describe the philosophy of tracking first. We suppose that all targets are known by the computer and that each target has its "inner flight number." Let us consider the coordinates of one single aerial rotation only. If we assume that nothing has been stored beforehand, the uncertainty of a coordinate's being the coordinate of a target is maximal. Let us assume we had the coordinates of all targets of the revolution T. Now the uncertainty about the positions of the targets during the revolution $T + 1$ is much smaller, as all targets must be contained in circle-shaped areas, the centres of which are the stored coordinates of the targets during the revolution T, the radius of each being determined by the maximal velocity allowed by the programme (this maximum is a parameter which can be set by the operator of the system). Let us now assume that we do know the coordinates found during the aerial rotations T and $T + 1$ of each target. The uncertainties about the positions of the targets are now even smaller still. The area resulting from the target's limited ability of motion is now smaller than the circle-area found from one point. We shall call this limited area the "expectation area." The number of former points, the coordinates of which have been incorporated in determining an expectation area, determines as in a Marcov algorithm its "degree." Figure 4 shows an expectation area of first degree being determined by the two positions P and Q. The next expectation area is being determined from the points Q and S.

The philosophy described in the paper uses expectation areas of second degree. There are always two consecutive "old" coordinates necessary for determining an expectation area.

We watch a single track: from the revolutions T and $T + 1$ we know the coordinates of the points P and Q. From these we can calculate an expectation area for the revolution $T + 2$. If a point S during the revolution $T + 2$ is a "hit" in this expectation area this coordinate is tried as a coordinate of the target (Fig. 4). From Q and S, then, the expectation area for the next revolution is found. As long as each expectation area receives just one hit tracking is rather simple. If there is no hit in the expectation area (e.g. if the coordinate has not been recognized by the digital detector) there are two possibilities: the algorithm of tracking may be cut off; or the missing

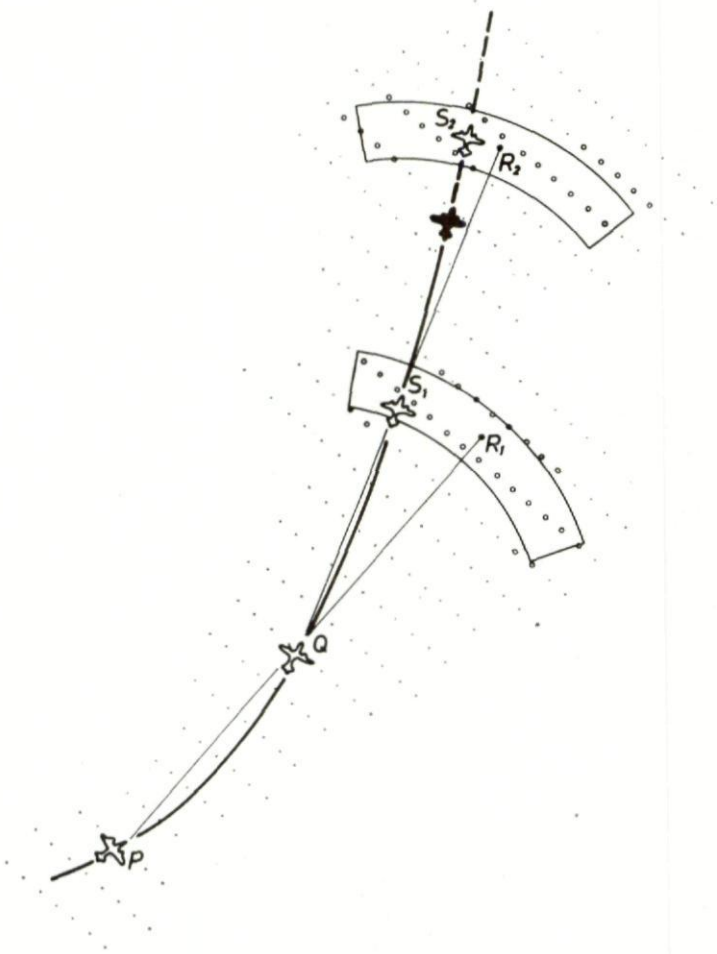


FIG. 4. *Expectation area calculated from the consecutive positions P and Q.*

signal is "bridged" by the setting up of a greater expectation area for the next revolution. An expectation area which bridges $n - 1$ missing signals is said to have the "rank" n. On the other hand it is possible for an expectation area to receive several hits. Not knowing which hit is the "true" one, each hit is used for setting up a track; thus "parasites" grow. We distinguish between such parasites being absorbed during further tracking and those diverging from the true one. Figure 5 shows the absorption of a parasite by the true track.

It is sufficient to distinguish between the following three "states" in an expectation area:

(1) The state "empty"—the expectation area has received no hit.

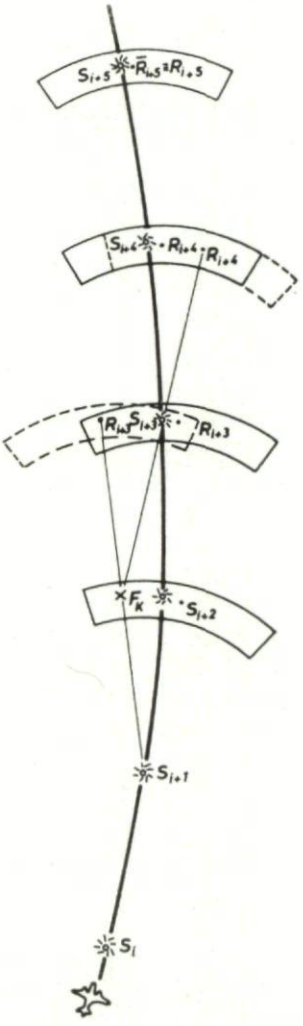


FIG. 5. *Absorbing of a parasite.*

(2) The state "critical"—the expectation area has received $k(v)$ hits or more, the "critical number" $k(v)$ depending on the rank v of the expectation area.

(3) The state "i hits" with $1 \leq i < k(v)$

The philosophy of tracking follows a very simple concept: The computer always has only got to know the state and the rank of an expectation area for it to decide on its future. In a tracking system of the rank 5 (that is a system using expectation areas of the rank up to 5) this concept can be formulated thus: an expectation area of the rank v which is found to be of the state

"empty" or "critical" is bridged by an expectation area of the rank $v + 1$ if $v < 5$. In the case of $v = 5$ the track is erased. An expectation area of rank v $(1 \leq v \leq 2)$ with i hits $(1 \leq i < k(v))$ causes the construction of i expectation areas of the rank 1, in the case of $3 \leq v \leq 5$ it causes the construction of i expectation areas of the rank 2.

This treatment results in the possible construction of i tracks, only one of which might be the true one. In a system with the rank 5 the computer is capable of bridging 4 missing signals. Thus parasites are dying off after a minimum of 5 revolutions unless they happen to receive a hit. This means that neither the computer nor the operator know which track is the right one during 5 aerial rotations. This difficulty is met by a modified tracking system which allows divergent parasites to die off with a rank lower than 5.

Let us assume that because of two hits in one expectation area a parasite has developed in addition to the true track. This parasite receives the same inner flight number as the true one, and that results in "related" expectation areas having the same "name." This parasite would not be dying off before 5 revolutions even if it didn't receive a further hit. We don't know which expectation area bears the true track and which bears the parasite. The erasing of a parasite requires a decision. The modified tracking programme contains a decision algortihm which always takes over control when there are "opposing" expectation areas which are pretending to be the true ones.

We shall call two expectation areas with the same flight number "opposed' if they both have the state "empty" or "critical." Let m be the rank in which a parasite is to be erased. The decision algorithm says:

An expectation area of the rank m, being "empty" or "critical" is to be erased, when there exists at least one expectation area of the same flight number with i hits, $1 \leq i < k(v)$.

An expectation area of the rank m is erased when it is opposed to one of rank smaller than m. In all other cases the expectation area of the rank is not erased.

The algorithm doesn't exclude the possibility of a false decision of erasing the true and letting the parasite survive.

The described tracking philosophy has an upper limit of stability; this is the number of signals being allowed for each aerial rotation. The system is growing unstable when the number of coordinates is so large that parasitic branches are enabled to grow forth. The stability limit depends on the rank used for the tracking: increasing the rank results in decreasing the stability limit. A low rank causes a more "nervous" system than a high one. Research has shown that the highest rank possible is the best for tracking, but a natural limit is imposed by the decrease of the stability limit: above a certain rank the limit would be lower than the number of targets to be expected, thus rendering tracking impossible.

A first test with a system of the rank 3 was run in Erbach (Ulm), rendering very reasonable results. In a second test to be run at Frankfurt we shall use higher ranks. Tracking assumes a one-time recognition of all aircraft by the computer. As new aircraft continuously demand new recognition as well as lost old ones, a permanent acquisition parallel to tracking is necessary.

The philosophy of acquisition is based on expectation areas of first degree. For the first recognition all coordinates of the first aerial rotation which are marked as "strong" coordinates by the digital detector are stored by the computer. As previously explained, each coordinate possesses a circle-shaped expectation area of first degree. If no coordinate of the second aerial rotation is found in this expectation area the point of the first rotation is erased.

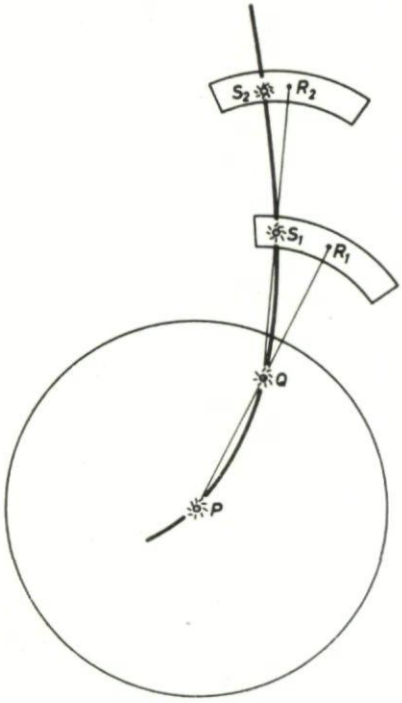


FIG. 6. *Acquisition of a target.*

If a target is found in this area the coordinates of these two points, the one of the first and the other one of the second revolution, are used to determine a second degree expectation area. Figure 6 shows this process of target acquisition. This new expectation area is erased if it doesn't receive a hit during the next revolution. If it does receive a hit it is called "target-suspicious" and it is handed over to the tracking-programme. This part of the surveillance-programme decides on the further evaluation of ths new track. We might say: a track is not called "target-suspicious" unless three successive signals simulate a track of an aircraft.

This acquisition programme is run permanently parallel to the tracking and is fed by those signals not being identified as hits in old expectation areas by the computer. Tracking is done with priority. The priority-regulation uses a two-level time-reserve-criterion. In a stationary run, control is at first executed by the tracking programme. After handling all signals, control is

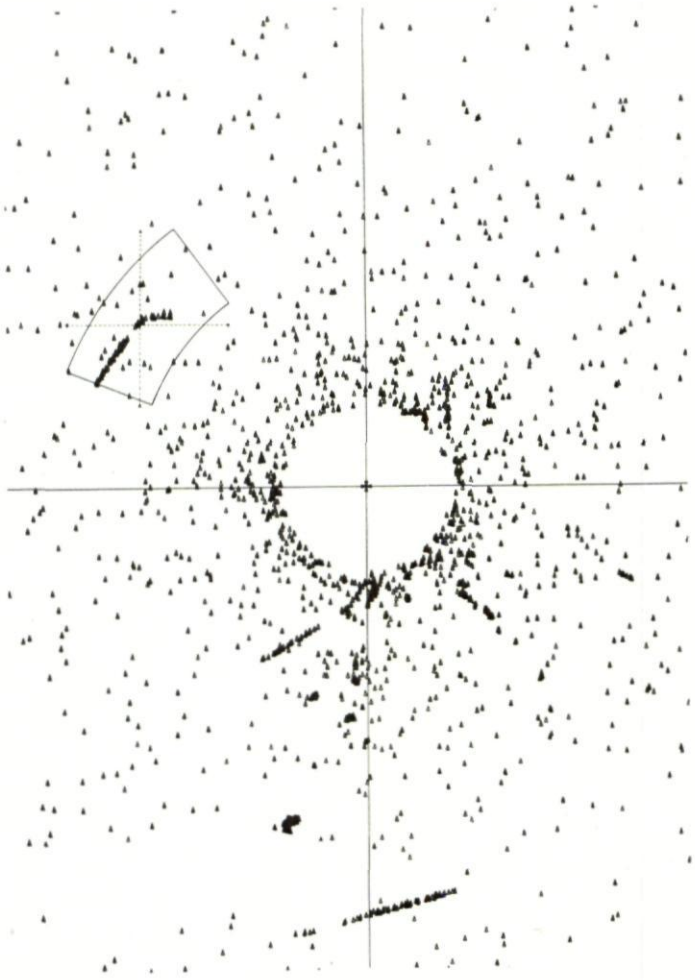


FIG. 7. *Detector outputs of several aerial rotations (as watched in Erbach) drawn by a plotter. (Graphomat.)*

handed over to the acquisition-programme. If the time-reserve is "1," that is "little time," the actual coordinates are tested as to whether they might be registered as hits in expectation areas of first degree resulting from the previous revolution. The hits are stored. Time-reserve "2," "plenty of time," causes the processing of such hits in expectation areas which were found in the previous revolution: expectation areas of second order are set up and handed over to the tracking programme. As long as time-reserve is "2," acquisition can be made.

The checking of the time-reserve is done for two reasons: on the one hand the computer is of limited capacity which cannot be outrun without upsetting

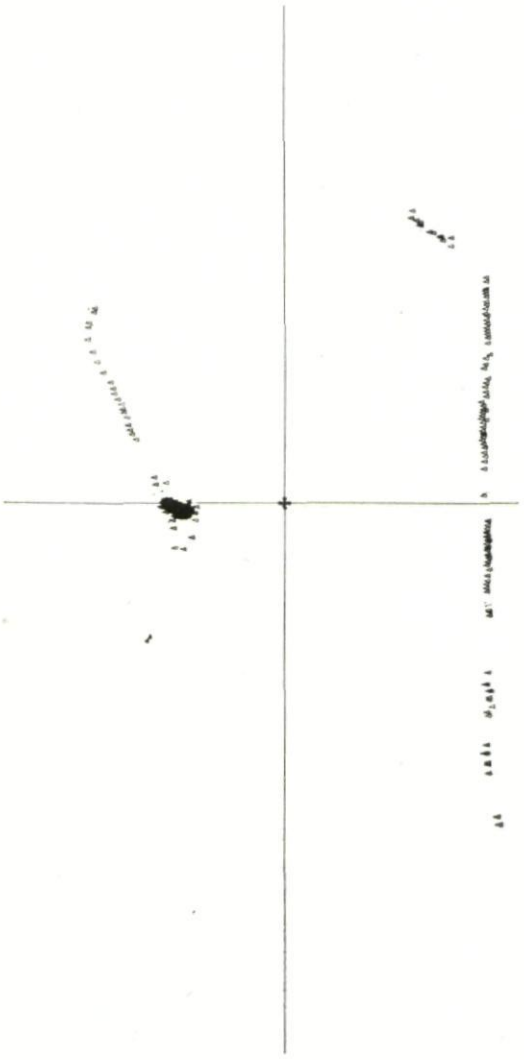


FIG. 8. *Signals of two aircraft tracked by the surveillance-programme drawn by a plotter. (Graphomat)*

the whole system; on the other hand computing time is to be used with optimal efficiency. The number of output data of the detector as well as the computer's processing time, are random numbers. Therefore the described organisation works as a "self-regulating system." If all the targets are not acquired during the first two aerial rotations, they will be acquired later on according to the capacity of the computer used.

These clutter-coordinates passing the MTI-threshold and being handed over to the surveillance-programme by the detector may unnecessarily abuse the computer's capacity by causing acquisition and tracking before recognition as clutter-signals. Thus we try to remove those coordinates situated in a clutter-area in advance. The easiest way of doing this is for the operator to enter the shape and the location of the clutter areas into the computer from outside. Yet, because of the computer's ability to recognize clutter signals after several rotations by comparing the computed speed of the target with a lower limit of speed, it is possible to lay out an algorithm enabling the computer to learn the size, shape and the location of such clutter areas. The same algorithm also enables the computer to forget about those areas under certain conditions, thus taking account of the fact that clutter areas may change during the run of the day.

The total processing system has several free parameters which during test-runs can be regulated manually providing means for statistical investigations. These parameters are: the rank of tracking, the rank of erasing, and the critical number of hits. It is planned to have these parameters regulated by the programme thus reaching a self-organising system. The described surveillance system is being tested in Frankfurt during the next weeks.

During our first test runs in Erbach using a surveillance-radar-system of Telefunken we studied the way the surveillance-system works. The output of the detector was a punched tape evaluated afterwards by a computer. We learned that automatic acquisition and tracking showed very reasonable results. A number of statistical investigations was made. For example, we were interested in the frequency of the use made of an expectation area of a certain rank by the system.

CHAPTER 29

EVALUATION OF TRACK-WHILE-SCAN COMPUTER LOGICS

L. LETH-ESPENSEN
SHAPE, Technical Centre, The Hague, Holland

The increasing number of vehicles which utilize the airspace demands a detailed knowledge of the positions, vectorial velocity, height and identification of all flying objects. This requirement is by no means just a military one, but pertains as well to the solution of the air traffic control problem. Automatic tracking of aircraft based on radar-derived data is a well-established method in "Airspace Management."

At the SHAPE Technical Centre the problem of the evaluation of the tracking performance of various proposed systems for airspace management was recognized several years ago. A number of methods have been tried, and experience gained accordingly. This paper will not discuss the whole background, but will highlight the important points which led to the method finally adopted.

1. TRACKING TERMINOLOGY

To set the stage for the following discussion it is necessary to explain briefly tracking methods and terminology. In this context the radar data will consist of positional data of flying objects as detected by a scanning radar with an associated data extractor. Each set of coordinates delivered to the tracker will constitute a so-called plot. Plots may not necessarily come only from flying objects, but may also be generated by noise and clutter. Furthermore, since the detection process is a statistical one, aircraft within radar cover may not necessarily be detected at every revolution of the antenna. The quality called the "blip-to-scan ratio" expresses the probability of extracting a plot from an object during a single scan.

In practice, owing to system noise and inaccuracies, the plots obtained from the radar are not true representations of aircraft positions. In general, the data will be corrupted by small errors in both azimuth and range.

We have established the necessary information to proceed with the tracking discussion. When the first plot is received, we are not able to start a track: at least two plots are necessary to determine a velocity vector and therefore to initiate a track. Under ideal circumstances, with no clutter or noise plots (sometimes called false alarms), we can associate two successively received plots as belonging to one and the same aircraft. A velocity vector, determined by the position difference and time interval between the two

plots, can be established and we are in business. When new plots are received, we can continue to associate them in a similar way, and we have established a track, i.e. a continuing, coherent record of the movement of a single aircraft. If real life were so simple, this paper would not have been given today, so we will proceed to discuss the difficulties, which make tracking more complicated.

The inaccuracies of the individual plot positions have already been mentioned. Figure 1 shows an example of what may happen in a real situation.

Correction formulae: $\Delta\bar{x} = \bar{x}_{\text{plot}} = \bar{x}_{\text{track,forecast}}$

Position: $\bar{x}_{\text{track,new}} = \bar{x}_{\text{track,forecast}} + \alpha\Delta\bar{x}$

Velocity: $\bar{v}_{\text{track,new}} = \bar{v}_{\text{track,old}} + \beta\dfrac{\Delta\bar{x}}{T}$

T: Time since last correction

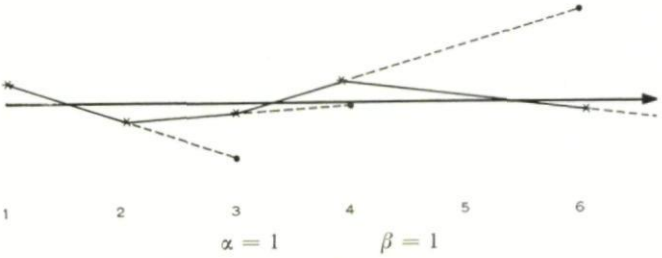


FIG. 1. *Simple track prediction.*

The thick line illustrates the actual aircraft path and the asterisks the plots received. Plots 1 and 2 enable a track to be initiated with an assumed velocity shown by the dotted line. This process is repeated. Notice that plot 5 is missing due to a fade and a correction cannot be made before plot 6 is received. The velocity vector is wagging like the tail of a happy dog and missing returns give rise to appreciable differences between estimated positions and the real flight paths.

To improve these matters a more complicated logic than the one just described must be introduced. In tracking terminology we talk about the smoothing of the position and velocity. Figure 2 shows what we would call a classical tracking logic, introducing the position and velocity smoothing parameters often called α and β respectively. Notice also the variable T which implies time since last correlation, a necessary factor in the calculation of velocity. By choosing different values of α and β—a subject which will be treated in more detail later—different behaviour of the tracking performance can be obtained. Also shown in Fig. 2 is the track behaviour using $\alpha = \frac{3}{4}$ and $\beta = \frac{1}{4}$ for the same data as shown in Fig. 1. During the initiation of the track—that is when the first two plots only have been observed—no smoothing can be applied and we have to return to the original simple procedure of extrapolating the track by assigning a velocity vector to the track based on the two plots only. This is equivalent to using $\alpha = \beta = 1$ during the track initiation procedure. Whenever there is a failure to detect

a plot, we have no choice but to extrapolate the track, based on the stored estimate of velocity, until a new return is obtained.

Only difficulties due to data error and blip-to-scan ratio have been taken into account so far. There are other matters which also complicate the tracking procedure. In particular, we assumed that only one aircraft was present within radar cover. This, of course, is a trivial case. In practice, we will have several aircraft present at the same time and possibly also noise plots and clutter returns. These extra plots complicate the matter in the sense that incoming plots must be associated with tracks or plots already existing. This is a track-plot correlation or allocation problem: we must determine

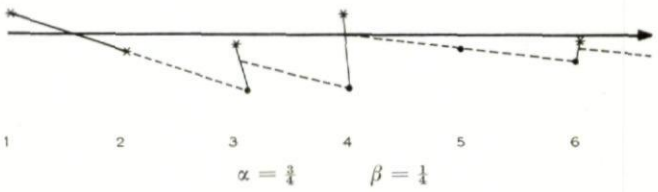


FIG. 2. *Prediction with position and velocity smoothing for this track* $\alpha = \frac{3}{4}$ *and* $\beta = \frac{1}{4}$.

which new plots belong to which tracks. Again, a classical method already exists which apparently can solve this problem: we introduce what is called a gate. Once a velocity has been established, we can calculate a forecast position of the track, that is, the position where we expect the next plot is to be found. The allocation process is then simply performed by associating a plot falling within a small area or volume, called a gate, placed around the forecast position of the track in question. Figure 3 shows an example. Two aircraft have established two tracks and the gates determine the areas within which we expect to find plots to associate with the two tracks. The dotted gates were used specifically for the track initiation procedure in exactly the same way. Once the allocation has been made, we can proceed to the necessary correction procedure. But we are still not quite happy! There are a few important cases left. What happens in particular when aircraft get close to one another? We will get overlapping gates, and we find plots which may be allocated to two different tracks! We could use smaller gates, but the gates must be large enough to encompass the correct plots. The plot position error, for example, sets a lower limit on the gate size even if the aircraft moves along a straight line. Furthermore, aircraft often turn—this complicates the matter further. Again, the classical solution consists of the introduction of several gates with different sizes: generally a so-called initiation gate, a small gate and a large gate. The size of the initiation gate is determined mainly by the maximum expected velocity of the aircraft to be tracked; the small gate by the data error and the large gate by the maximum acceleration of the aircraft during turns. Furthermore, considerations must be included which take into account the effect of missing returns due to blip-to-scan ratios less than one. Different smoothing parameters are generally associated with the different gate sizes.

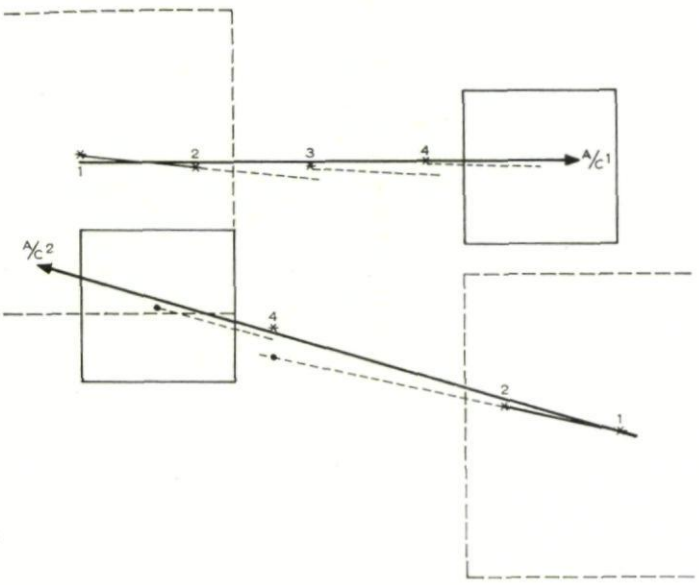


FIG. 3. *The use of gates.*

One thing still needs explanation, namely, the criteria for determining that a track is terminated. Some tracking logics carry a so-called "track quality number" which is increased up to a maximum value with different increments, depending on the type of correlation found (initiation, small gate, large gate) and decreased when no correlation is found. When the quality number reaches a certain lower limit the tracker considers the track as dead and terminates it.

2. THE EVALUATION METHOD

The design of a general method for tracking evaluation requires various concepts to be defined. These decisions may significantly influence the final results of the evaluation. Therefore, even though agreement may exist on which are the major tasks of a tracking system, this does not immediately solve the question of the most suitable procedures, or criteria for tracking evaluation. Some of the difficulties involved in the development of a method of evaluation which is both comprehensive and objective are due to the scarcity of objective criteria on which to build measures of effectiveness. These measures of effectiveness must be such that they allow both an absolute evaluation of the capability of a particular system and a comparison between the performance of different tracking logics. In other words, any result of the tracking evaluation is influenced by the decision criteria. These must therefore be relevant to the objectives of tracking, and be fair, complete, and quantifiable.

In a tracking system, the design of a set of decision criteria also constitutes an important problem. These decisions concern the acceptance or rejection of a plot as belonging to a particular track. Wrong answers to this question cause missing correlations, or tracks to be maintained on a sequence of accepted plots which are thought to originate from the same target but which, in fact, belong to different aircraft. The ability to make correct decisions in allocating incoming plots to existing tracks or initiation of new tracks is an essential factor and reflects very much the capability of a tracking system. This is consequently of importance for the design of the tracking evaluation method. To handle this problem, it was found necessary to define two types of tracks, the Apparent track and the Real track. A logical process involving the correlation error, in addition to missed plots, may be used also. In principle, any process for track termination which permits a few scans of extrapolation, but prohibits persistence of tracks for many scans without input data may be considered "reasonable", tempered by the requirements of the users.

An automatic tracker will normally enable many aircraft to be tracked simultaneously and information on each track will be contained in separate computer locations of the tracker, generally referred to as individual track "slots," or "channels." The track slot number will in this context refer to an internal location, or address given by the tracker to any individual track carried by the system.

These examples were given to establish the terminology, but as you will hear later, there may be many alternative methods of performing the basic functions of a tracking system. To review, these basic functions are:

Initiation

Allocation

Correction and Prediction

Termination

The crux of the matter is that although everything seems so easily solved from the outset, experience shows that several of the criteria which we can establish for improving tracking ability in different situations are directly contradictory. Large gates would, for example, enable better ability to follow turns but would obviously lead to difficulties in the allocation area when the density of aircraft gets high.

In the ideal tracker, aircraft are tracked in one and the same track slot from the time when they enter radar coverage until they leave it again. In practice it does not necessarily work that way. Tracks die and tracks switch between different aircraft.

We will go into further explanations of these phenomena later; for the moment we will just assert that they occur, and state that it is quite possible for an automatic tracking system to produce an output which bears no resemblance to the real situation.

3. DEFINITIONS

An Apparent track is a track as it appears to the Supervisor of the Tracking System. It starts when the track is initiated and ends at the last scan for which correlation has been obtained. To make clear what is meant by an Apparent track, an example is given in Fig. 4. The vertical division of the two rows into squares indicates a sequence of consecutive scans. The first row contains scan numbers, and the numbers in the second row indicate the type of correlation which the track has obtained.

The duration of the Apparent track can be found from the string of correlation numbers given in the example. The correlation row shows a "2"

APPARENT TRACK

Example of track slot history

Scan Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Type of Correlation	−1	1	2	3	4	3	0	4	4	3	4	4	0	0	0	0	0	0	−1

Duration: from initiation as determined by tracker to last correlation before termination

Correlation Legend	Types of Correlation
−1	Track slot is empty
0	No correlation in present scan
1	Pending plot, in this example first of two plots necessary for track initiation
2	Track initiation
3	Large gate correlation
4	Small gate correlation

(Minimum duration of an apparent track necessary for acceptance (MA) = 3).

FIG. 4. *Slot history for an apparent track.*

in the third scan, indicating track initiation. This is the start of the Apparent track. The development of the track is expressed through the correlation numbers of the following scans. The sequence of correlations is: Large gate, Small gate, Large gate, No correlation, Small gate, and so on. After several scans, the internal termination logic decided that the track was lost and cleared the track slot. On the 19th scan, the type of correlation is set to minus one, indicating that the track is terminated, and the slot is free and available for the development of a new track. From the definition of an Apparent track, it can now be decided that scan 12 represents the end of the track and consequently that the track duration is ten scans. A precise definition of the beginning and the end of an Apparent track in this example is the following:

First scan : Scan of track initiation, i.e. type of correlation equal to two

Last scan : Latest scan with a correlating plot, i.e. the type of correlation different from zero, ahead of scan of track termination.

It is immaterial for an Apparent track whether the string of correlating plots originates from one aircraft or from several, or whether they are false returns. This question is a primary one, however, for the Real track.

A Real track is the part of an Apparent track in which the same aircraft has been tracked. Consequently, several Real tracks may occur within an Apparent track. If the track logic in use permits correlation with more than one plot per scan, overlapping Real tracks may be experienced.

The correlations with a certain aircraft constituting a Real track do not need to be consecutive. The aircraft is considered tracked even if, for some intermediate scans, the plots from the aircraft do not correlate with the track.

REAL TRACK

Extended example of track slot history

Scan Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Type of Correlation	−1	1	2	3	4	3	0	4	4	3	4	4	0	0	0	0	0	0	−1
Aircraft Number*	8		0	8	8	0		9	8	8	10	10							

1.Real a/c 8 — 2.Real a/c 10

←apparent track→

Initiation: Minimum number of correlations required for acceptance of real track (MS= 2)

Termination: Maximum number of consecutive non-correlations permitted (MI) = 4

Duration: From first plot of the particular aircraft leading to initiation to last correlation for particular aircraft before termination

* Aircraft Number = 0 means a noise plot.
Notice that the aircraft number is not available to the tracker.

FIG. 5. *Slot history for a real track.*

For this reason, the concept of a Real track as defined in the present method of evaluation, involves two parameters. One of these parameters gives a value of the minimum number of correlations (with the same aircraft) required to constitute a Real track (MS). The other sets the maximum number of consecutive missing correlations (with respect to the particular aircraft) permitted within a sequence of correlations which in other respects fulfils the requirements for a Real track (MI).

Figure 5 shows an example of how the Real tracks are computed. This is a repeat of Fig. 4, to which has been added a third row, showing the aircraft number for the correlating plot. Looking at the third row, it appears that the plot obtained in scan 2 belongs to aircraft No. 8. Furthermore, the initiation in the third scan was made on a noise plot as indicated by an aircraft number equal to zero. Proceeding in this way, scan by scan, the aircraft numbers on which the tracks are constructed can be mapped. Knowing these numbers, the Real tracks belonging to this Apparent track

may easily be computed. However, prior to these computations, the values of the two decision parameters must be fixed. In the present example the minimum number of correlations required (MS) is set equal to two and the maximum number of consecutive non-correlations (MI) is given the value four. The string of aircraft numbers shows that the first Real track starts at scan 4 and ends with scan 10. This track contains aircraft no. 8. The noise plot is not counted in the duration of the first Real track.

Since four consecutive non-correlations are permitted, the two correlations in scans 9 and 10 are considered as belonging to the Real track started in scan 4, i.e. constitute a part of the first Real track. It should be clear that a non-correlation means either that the expected plot has faded or that it was not allocated to the proper track. As MS = 2, the two correlations with aircraft No. 10 make a second Real track.

Several investigations have been made into the influence of these parameters, and it turns out that MS = 3 and MI = 5 lead to consistent answers in the evaluation. Further increase in the parameter values does not change the performance noticeably.

The concept of Real and Apparent tracks constitutes an important part of the evaluation criteria. However other properties of a tracking system also require attention, in particular, the ability to follow tracks successfully around turns. It has already been mentioned that straight-line tracking and turn-following impose contradictory requirements on the logic when the density of flying objects becomes high. The evaluation must therefore also take this fact into account.

4. PERFORMANCE CRITERIA

We might define specific situations for evaluation with constant aircraft densities and specific ratios of turns versus straight-line tracks. However such situations are bound to be arbitrary selections which may or may not bear much resemblance to real situations. Therefore optimizing the logic for such selected situations does not necessarily lead to the best solution for other cases. To avoid this difficulty it has been found useful to perform the evaluation based on data which constitute many different situations of varying complexity. In this way the comparison of different logics becomes relatively easy; the difference in performance is clearly expressed in the results. Admittedly no absolute value of tracking ability is found in this way, but such a measure does not make much sense anyway since one can always define situations which nearly any logic can cope with and others which will be impossible for even the best one.

To avoid misunderstandings at this point, let it be clearly stated that specific situations are always useful for the study of particular improvements of a logic, but the overall evaluation must eventually be based on the whole spectrum of situations, presented simultaneously. The non-trivial situations for tracking must therefore contain high densities of aircraft, crossing tracks and turns. Experience has shown that the difficulty of resolving track crossings is increased when the velocity vector difference between the tracks

is small, e.g. when the two tracks stay in close proximity for a long time. Slightly doglegged tracks (mainly straight-line tracks which change heading slightly from time to time) going in the same direction with nearly the same velocity produce such cases. Advantage has been taken of this fact in the generation of air situations. Furthermore, turns are introduced in such a way that they occur against a wide range of background densities.

Individual tracks are divided into three distinct parts, namely, before turn, in turn, and after turn. The aircraft density decreases slowly as time goes by.

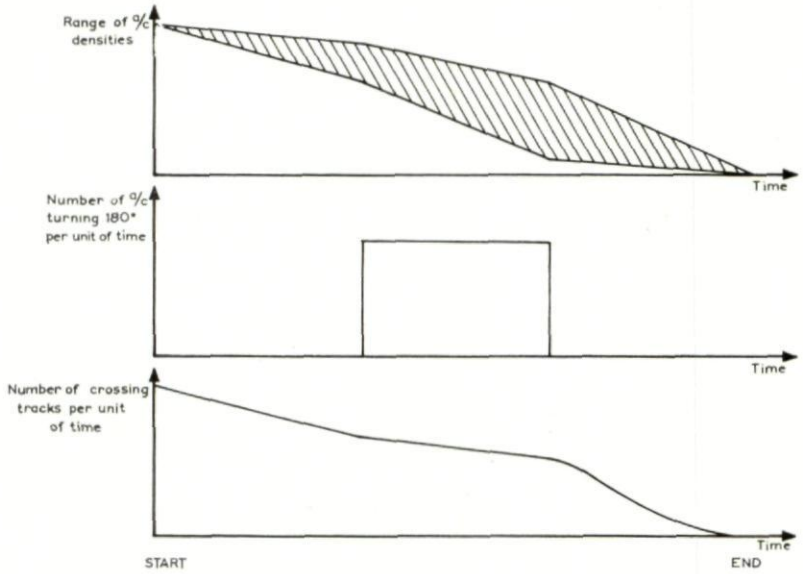


FIG. 6. *Input air situation behaviour.*

A typical input air situation behaviour is shown as a function of time in Fig. 6. It is important to understand the implication of this behaviour in connection with the figure of merit for the tracking performance derived later. In particular, notice should be taken of the fact that the density gradually decreases—easing the tracking difficulty—once the aircraft has finished its turn. This implies that a track followed around a turn has a relatively high probability of being tracked successfully thereafter. This fact is used as a weighting factor in the figure of merit for awarding turn-following capability to the track logic.

We can now turn back to a more detailed description of the actual evaluation. As already explained, tracks are divided into two major groups namely, Apparent tracks and Real tracks. Apparent tracks are again subdivided into the following classes.

(a) Apparent tracks with a duration shorter than a certain minimum number of scans determined by the parameter MA (MA = 3 is often used); these short tracks are not considered for further evaluation.

(b) Apparent tracks which contain no Real tracks according to the definitions specified by MI and MS. These tracks are called Empty Apparent tracks. They constitute initiations which are based on random jumping between neighbouring plots without ever producing any useful track. The better the initiation logic, the smaller the number of Empty Apparent tracks.

(c) All other Apparent tracks containing one or more Real tracks.

Real tracks are divided into two classes:

(d) Real tracks initiated before the turn.

(e) All other Real tracks.

Notice that Apparent tracks cannot originally be defined as initiated before or after the turn, since they may consist of Real tracks of which some have and some have not started the turn.

The mean value of the track duration expressed in scans is calculated for all Apparent tracks (longer than MA) and for Real tracks started before and after the turn.

A weighting factor is calculated as follows:

$$W = \frac{\text{Number of samples in } c}{\text{Number of samples in } (d + e + b)}$$

where b, c, d and e refer to the previous description of classes.

The product of the weighting factor and the mean value of Real tracks initiated before the turn is used as a figure of merit for the tracking system:

$$F = W\bar{d}$$

The figure of merit developed in this way is used for comparison of different logics using the same air situation as input. It is also useful for comparison of results obtained by varying the parameter values of, for example, data errors and blip-to-scan ratio in the same air situation.

In the ideal case W reaches the value 1 which indicates that all Apparent tracks contained exactly one Real track, and $\bar{d}$ reaches the value which is obtained when all aircraft are tracked from the very beginning to the very end.

It may at first look surprising that the number of turns survived does not enter the formula. However, this point is taken care of automatically. Two specific examples will explain this mechanism in principle. Suppose we have two track logics, of which one is absolutely ideal in all respects whilst the other suffers in the sense that all tracks die when turning. In both cases all Apparent tracks will contain exactly one Real track and there will be no Empty Apparent tracks. This means that W in both cases reaches the value 1. However in the ideal tracker $\bar{d}$ will reach its maximum value $\bar{d}^*$ and the number of samples in class e will be zero, whilst in the tracker in which no tracks survive the turn the number of samples in class e will be the same as in class d, namely, equal to the total number of aircraft, and in this case $\bar{d}$ will be slightly less than $\frac{1}{2}\bar{d}^*$. The reason is that each aircraft path has been

divided by the logic into two separate tracks, namely, before and after the turn.

Another case could also be imagined, namely, the one where aircraft turns resulted in track switching in such a way that each Apparent track contained two Real tracks. In this situation the weighting factor W would be halved, penalizing the figure of merit for carrying tracks with wrong identification.

Empty Apparent tracks cause trouble by injecting misinformation into the system. They are introduced into the calculation of W in a manner which penalizes this behaviour.

It may be argued that the chosen figure of merit is a rather arbitrary measure of the tracking performance. This is true to the extent that two systems with nearly the same resulting figure of merit must be compared in more detail to explain what the particular difference between them may be. However, if two systems show a marked difference in the figure of merit, we generally have a clear picture of which system is the better.

Later, when we turn back to a discussion of the improvements obtained by various logics the results will be presented in a way which contains all the information available from the evaluation procedure, rather than just the figure of merit.

5. EVALUATION PROCEDURE

Until now the discussion has been concerned with definitions for evaluation. How it is actually done is the subject of this section. It was pointed out previously that the actual identification of the individual aircraft will not in general be available to the tracking system and that this was nevertheless one of the necessary data used in the evaluation for the determination of Real versus Apparent tracks. To avoid this practical difficulty in real life, we have based our method on a completely controlled experiment performed by means of digital computer simulation.

SUMMARY OF DATA USED IN EVALUATION

Apparent tracks: defined by track logic and MA

a: Apparent tracks with a duration shorter then MA (described)

b: Apparent tracks containing no Real tracks (Empty Apparent tracks)

c: Apparent tracks containing at least one Real track

Real tracks: defined by MI as MS

d: Real tracks initiated before turn

e: Real tracks not initiated before turn

$$\text{Weighting Factor} \quad W = \frac{c}{d + e + b}; \quad W \leq 1$$

b, c, d and e refer to the number of samples found in each class.

Figure of merit: $F = W = \bar{d}$

$\bar{d}$ is the mean track life expressed in scans for Real tracks initiated before turn.

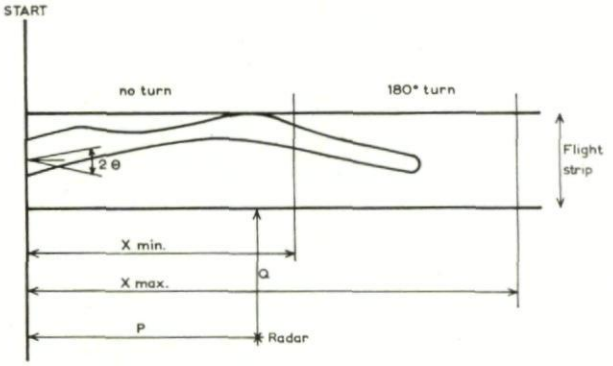
FIG. 7. *Summary of data used in evaluation.*

The tracking procedure obviously suggests computer simulation, since the behaviour of one computer can be simulated uniquely on another. This part of the investigation does not therefore give rise to many problems.

We must however create the air situation which constitutes the input to the simulated tracker.

6. GENERATION OF AN AIR SITUATION

The aircraft path which was wanted as input has already been briefly described. To generate the simulated radar data corresponding to such situations a computer program was written. Figure 8 shows the main



N	:	Number of aircraft
D	:	Density of aircraft at start
Sp min., Sp max.	:	Minimum and maximum speed of aircraft
H min., H max.	:	Minimum and maximum altitude of aircraft
X min., X max.	:	Minimum and maximum distance flown before 180° turn
FSW	:	Flight strip width
P and Q	:	Position of radar relative to start line and flight strip
DR	:	Data rate (radar rotation rate)
RT	:	Rate of turn for 180° turn
θ	:	Deviation interval from main flight direction at start
False alarm	:	False plot density due to noise
Markov	:	Markov matrix for transition probabilities of heading changes
Unit	:	Unit change in heading due to Markov process or leaving flight strip

FIG. 8. *Basic input for air situation generator.*

features of this program. The basic input data are more or less self-explanatory. The individual aircraft paths are developed on a scan-to-scan basis, and the intersections of the radar beam with the paths are calculated for every rotation of the radar. The main direction of the flight paths is that of the defined flight strip, but the initial heading of any individual aircraft may deviate from this direction at start time with a maximum value defined by θ. A Markov matrix giving transition probabilities for different states of heading changes is also introduced. The different states of the matrix define heading changes compared with the one defined at the starting time. The resulting

aircraft paths will in this way consist of straight-line segments with slight changes of heading from time to time.

Over and above this process there is another one which keeps the aircraft reasonably within the flight strip. When aircraft fly out of the defined area they are slowly turned back in a way which roughly mirrors their path in the flight strip border. This procedure is necessary to avoid a complete spread of the tracks as time goes by.

Notice that each aircraft performs one 180° turn with a specified rate of turn. After turning, aircraft fly back in the same way as they came in and the flight path is terminated when the start line is reached again.

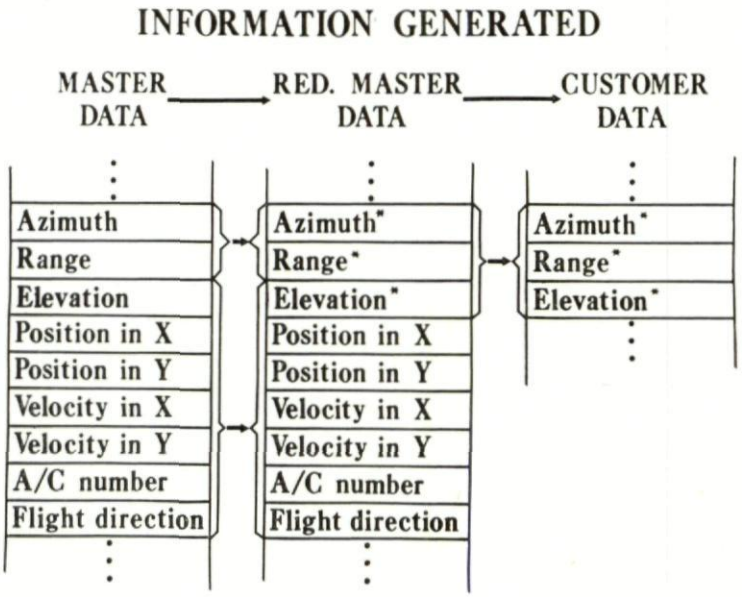


FIG. 9. *Data processing.*

The output from the air-situation generator consists of data as they would have been seen by an ideal radar with a perfect identification system. For each scan pertinent information on all plots is delivered in sequential order in azimuth on a magnetic tape, the so-called Master Tape. False plots can also be generated and delivered as a part of the output sorted among the plots belonging to the aircraft.

Figure 9 shows the information content on the master tape for an individual plot. The master tape is not used directly as input. A second program called the "disturber program" is introduced. This program is used for the generation of two tapes, namely, what we call the "reduced master tape" and the "customer tape." The point is that to simulate various realistic situations, data errors and blip-to-scan ratio must be introduced. The disturber program requires, apart from the master tape, input data which define standard deviations in range and azimuth (possibly also in height) and

a blip-to-scan ratio. A noise-reduction factor is also introduced which may decrease the number of noise plots on the master tape with any desired ratio.

Figure 9 indicates with an asterisk which data have been disturbed compared with the master tape. Furthermore a certain proportion of the plots has been removed depending on the defined blip-to-scan ratio. The reduced master tape contains both the disturbed data and the auxiliary and correct data for the plots which have been left over in the blip-to-scan ratio process. The customer tape only contains the disturbed data, corresponding to the typical output of a scanning radar plus extractor system.

The customer tape is used as input for the tracking simulation program; rigid computer format specifications have been laid down to ensure compatibility between input and program.

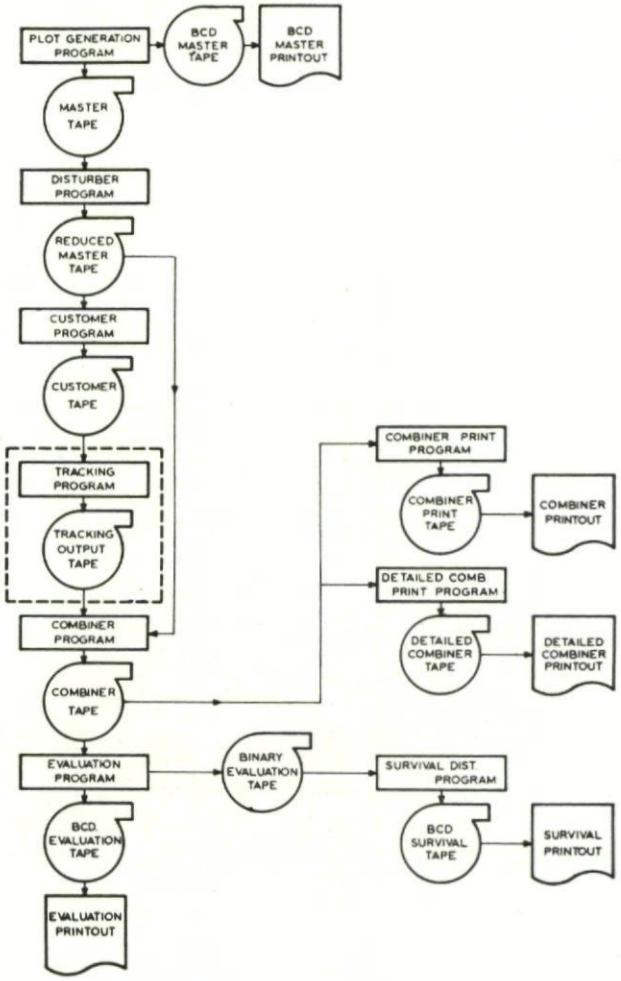


FIG. 10. *General program layout.*

7. PROGRAM SEQUENCE

Figure 10 illustrates the sequence of procedures used in the evaluation. The master, reduced-master and customer-tape generation has already been described. Next the tracking simulation is performed. What the tracking simulation program does to the data is, in principle, immaterial, as long as it accepts the input data and delivers the output data in the format specified by the evaluation system.

The figure shows that the tracker output and the reduced master tape are

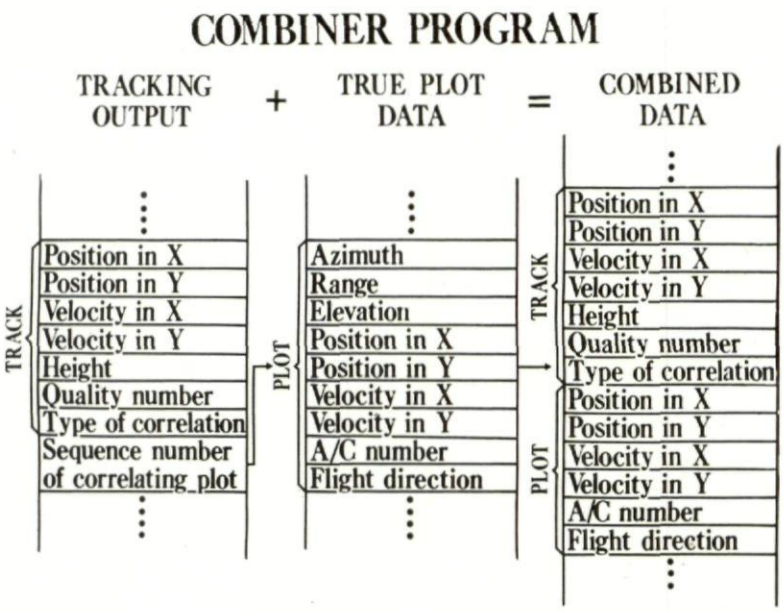


FIG. 11. *Data combining program.*

used for the generation of what is called the "combiner tape." This process is the key to the whole evaluation. The combiner program associates the information produced by the tracker with the exact information existing on the reduced master tape. Figure 11 shows the process performed. Notice how we have now produced the information needed for the evaluation of real tracks. Also available on the combiner tape are data such as track versus aircraft position and velocity. This information is not used directly in the evaluation process but it may be useful for several other purposes.

A combiner print program produces the kind of information which was used for the explanation in connection with the determination of Apparent tracks. An example of such a print is shown in Fig. 12. This presentation is extremely useful for finding logical errors in the track simulation program. The combiner tape is used for the generation of two other outputs, namely, the track break-down output and the survival distribution generator.

The track break-down program output is shown in Fig. 13. This print

NCODEC 81005 NSEQO 1

SHEET NR 7

Slot scans	73	74	75	76	77	78	79	80	81	82	83	84	NDROP
0.	011 00	011 00	011 00	011 00	011 00	011 00	211 00	011 00	411 00	111 00	511 00	011 00	0
1.	009 00	009 00	009 00	009 00	009 00	009 00	812 00	009 00	412 00	512 00	512 00	009 00	0
2.	011 00	011 00	011 00	011 00	009 00	611 00	913 00	211 00	413 00	113 00	513 00	1011 00	0
3.	009 00	009 00	012 00	009 00	1111 00	612 00	002 00	612 00	002 00	114 00	514 00	1012 00	0
4.	011 00	009 00	009 00	009 00	1112 00	613 00	001 00	009 00	001 00	003 00	115 00	1013 00	0
5.	009 00	1211 00	1311 00	011 00	1113 00	614 00	000 00	011 00	000 00	114 00	516 00	1014 00	0
6.	1412 00	009 00	1312 00	009 00	1114 00	615 00	009 00	009 00	009 00	003 00	521 00	415 00	0
7.	1413 00	211 00	1313 00	1011 00	003 00	616 00	1211 00	011 00	011 00	002 00	001 00	416 00	0
8.	1414 00	911 00	002 00	009 00	1114 00	622 00	1512 00	009 00	009 00	009 00	533 00	423 00	0
9.	1415 00	009 00	001 00	711 00	1115 00	002 00	1513 00	1611 00	1911 00	111 00	003 00	441 00	0
10.	004 00	1411 00	000 00	2111 00	1123 00	633 00	1514 00	1712 00	1912 00	111 00	533 00	001 00	0
11.	1415 00	2212 00	1311 00	1711 00	1133 00	633 00	1515 00	2113 00	2713 00	2411 00	533 00	433 00	0
12.	1423 00	009 00	1312 00	2012 00	1133 00	643 00	1516 00	2814 00	2714 00	009 00	534 00	441 00	0
13.	1433 00	2412 00	1313 00	2313 00	003 00	634 00	1523 00	3115 00	2715 00	2412 00	531 00	433 00	0
14.	1433 00	009 00	1322 00	3214 00	003 00	642 00	1541 00	004 00	2716 00	2913 00	533 00	433 00	0
15.	003 00	011 00	1333 00	3415 00	1133 00	633 00	1533 00	003 00	2723 00	2914 00	533 00	431 00	0
16.	003 00	009 00	1341 00	3616 00	1133 00	633 00	003 00	002 00	2731 00	2915 00	003 00	001 00	0
17.	1431 00	1911 00	1340 00	005 00	1134 00	641 00	1533 00	009 00	3033 00	2916 00	542 00	442 00	0
18.	1443 00	009 00	000 00	004 00	1133 00	643 00	003 00	4111 00	2432 00	005 00	002 00	434 00	0
19.	003 00	009 00	1340 00	003 00	1132 00	633 00	1633 00	4112 00	3034 00	2923 00	002 00	533 00	0
20.	1433 00	011 00	000 00	009 00	002 00	633 00	003 00	4113 00	3042 00	2933 00	002 00	433 00	0
21.	003 00	009 00	1333 00	2711 00	1134 00	003 00	003 00	4114 00	002 00	2943 00	002 00	441 00	0
22.	1433 00	4111 00	1333 00	1912 00	1134 00	642 00	003 00	003 00	3033 00	2934 00	002 00	440 00	0
23.	1434 00	009 00	1333 00	2413 00	1133 00	642 00	003 00	4114 00	2633 00	3055 00	534 00	433 00	0
24.	1433 00	2711 00	1333 00	002 00	1141 00	641 00	003 00	4115 00	3041 00	2634 00	535 00	444 00	0
25.	1442 00	2912 00	1333 00	1913 00	001 00	633 00	003 00	4122 00	3033 00	2643 00	541 00	435 00	0
26.	1433 00	009 00	1341 00	002 00	1133 00	633 00	009 00	4133 00	3041 00	2634 00	533 00	434 00	0

FIG 12. *Combiner print program.*

NCODEC 81005 NSEQO 1

SHEET NR 7

Slot scans	73	74	75	76	77	78	79	80	81	82	83	84	NDROP
27.	003 00	4911 00	533 00	1913 00	003 00	633 00	4411 00	4142 00	3042 00	2633 00	1331 00	435 00	0
28.	1443 00	5012 00	533 00	1914 00	1133 00	641 00	009 00	4134 00	002 00	003 00	001 00	435 00	0
29.	1433 00	009 00	533 00	3523 00	003 00	633 00	4411 00	4133 00	002 00	2642 00	001 00	434 00	0
30.	1443 00	011 00	533 00	3533 00	1131 00	643 00	009 00	4133 00	2934 00	2634 00	1540 00	435 00	0
31.	1443 00	009 00	1334 00	3533 00	1133 00	933 00	009 00	4133 00	2941 00	2633 00	000 00	434 00	0
32.	1434 00	009 00	634 00	3532 00	1142 00	933 00	4912 00	4133 00	001 00	2634 00	000 00	443 00	0
33.	1433 00	5011 00	532 00	3533 00	1133 00	633 00	009 00	4133 00	2933 00	2642 00	000 00	434 00	0
34.	1442 00	009 00	634 00	003 00	1143 00	931 00	009 00	4134 00	2941 00	2634 00	000 00	433 00	0
35.	1443 00	5011 00	535 00	1934 00	1143 00	1340 00	011 00	4133 00	2933 00	2733 00	000 00	434 00	0
36.	1444 00	5012 00	531 00	1931 00	1142 00	941 00	009 00	003 00	2934 00	2744 00	000 00	004 00	0
37.	1435 00	5013 00	530 00	1942 00	1134 00	943 00	009 00	4141 00	004 00	2632 00	009 00	004 00	0
38.	005 00	5014 00	531 00	1933 00	1133 00	933 00	009 00	4133 00	2941 00	2741 00	2611 00	1432 00	0
39.	005 00	5015 00	001 00	1933 00	003 00	942 00	009 00	4134 00	2940 00	2743 00	3712 00	1434 00	0
40.	434 00	5016 00	533 00	003 00	1141 00	941 00	009 00	004 00	2933 00	2733 00	3713 00	1442 00	0
41.	1441 00	5022 00	534 00	1933 00	1133 00	943 00	011 00	4133 00	2942 00	2733 00	3714 00	002 00	0
42.	433 00	5034 00	542 00	3734 00	1142 00	943 00	009 00	4133 00	002 00	003 00	003 00	002 00	0
43.	431 00	004 00	542 00	3742 00	1143 00	942 00	011 00	4141 00	2934 00	2641 00	1914 00	002 00	0
44.	433 00	004 00	534 00	2631 00	1141 00	943 00	009 00	4530 00	2933 00	2733 00	3715 00	1433 00	0
45.	003 00	004 00	533 00	2643 00	1133 00	933 00	5011 00	4533 00	2933 00	003 00	3723 00	1441 00	0
46.	434 00	004 00	542 00	003 00	1133 00	942 00	5012 00	4533 00	003 00	2644 00	1933 00	1432 00	0
47.	444 00	004 00	542 00	003 00	003 00	002 00	5013 00	003 00	003 00	2642 00	1933 00	002 00	0
48.	004 00	004 00	540 00	003 00	003 00	942 00	5014 00	003 00	2933 00	3733 00	003 00	002 00	0
49.	004 00	009 00	541 00	003 00	1133 00	932 00	5015 00	003 00	2941 00	3633 00	2741 00	1433 00	0
50.	441 00	4111 00	540 00	003 00	1133 00	933 00	5016 00	003 00	001 00	3733 00	2732 00	1434 00	0
51.	442 00	4112 00	533 00	1934 00	1143 00	933 00	5021 00	003 00	2933 00	3734 00	2733 00	233 00	0
52.	442 00	4113 00	003 00	004 00	003 00	933 00	5043 00	003 00	003 00	3733 00	2633 00	003 00	0
53.	002 00	002 00	533 00	2755 00	1134 00	941 00	5033 00	009 00	2933 00	3733 00	2633 00	003 00	0

FIG. 12. *Combined print program.* (Contd.)

Apparent track no.	No. of real tracks	Track slot no.	Rem. turns	No. of cons. no.-corr.	MI 6 A/C no.	MS 7 Dir. Init.	MS 7 At end	MA 3 Scan of app.	Init. real	Duration app.	Duration real	XSL. at app.	XSL. Init. real	YSL. at app.	YSL. Init. real
1	0	79						13		7		172		40	
2	1	83						6		25		177		31	
			8	7	5	1	1		6		21		177		31
3	0	46						23		15		160		50	
4	1	62						21		22		174		34	
			8	7	37	1	1		21		18		174		34
6	1	80						25		22		173		30	
			8	7	41	1	1		25		19		173		30
7	0	60						29		19		154		38	
8	1	41						45		11		141		33	
			8	7	44	1	1		45		7		141		33
9	2	73						12		50		173		32	
			8	7	14	1	1		12		30		173		32
			8	7	4	1	1		40		19		117		32
10	1	28						43		25		114		48	
			6	7	6	1	0		43		23		114		48
11	2	75						14		55		168		35	
			8	7	13	1	1		14		18		168		35
			6	7	5	1	0		27		42		142		42
12	2	83						45		28		125		32	
			8	7	26	1	1		52		13		111		32
			8	7	27	1	1		65		7		88		37
13	1	74						65		12		102		34	
			8	7	44	1	1		65		12		102		34
14	1	57						18		75		165		42	
			7	7	15	1	0		25		68		149		44
15	2	58						31		69		166		33	
			8	7	45	1	1		31		58		166		33
			8	7	39	1	0		87		13		45		38
16	1	64						9		92		172		39	
			6	7	7	1	0		25		76		141		42

FIG. 13. *Track break-down program output.*

MI 6 MS 7 MA 3

Apparent track no.	No. of real tracks	Track slot no.	Rem. turns	No. of cons. no.-corr.	A/C no.	Dir. Init.	At end	Scan of app.	Init. real	Duration app.	Duration real	XSL. at app.	Init. real	YSL. at app.	Init. real
17	2	42						19		83		167		46	
			8	7	17	1	1		19		17		167		46
			6	7	25	1	0		32		70		145		45
18	1	53						35		69		154		51	
			8	7	40	1	1		35		69		154		51
19	4	61						9		98		169		45	
			8	7	2	1	1		9		7		169		45
			8	7	7	1	1		16		9		158		41
			8	7	2	1	1		25		28		141		41
			6	7	14	1	0		51		56		95		36
20	1	44						27		94		175		47	
			7	7	42	1	0		27		92		175		47
21	1	31						21		103		163		50	
			4	7	20	1	0		22		98		162		49
22	3	67						21		103		167		34	
			8	7	24	1	1		21		37		167		34
			8	7	25	1	1		50		17		109		44
			8	7	24	1	1		67		51		75		45
23	1	60						71		55		90		44	
			0	7	4	0	0		71		55		90		44
24	2	43						52		78		114		32	
			7	7	23	1	0		52		44		114		32
			8	7	40	1	1		104		23		24		41

FIG. 13. *Track break-down program output.* (Contd.)

DURATION OF APPARENT TRACKS

Minimum number of scans considered = 3
Number of empty apparent tracks = 5

$Z = 0.662$ $F = 45.60$

Number of observations = 93
Mean duration = 107.83

Duration more than (scans)	Relative frequency
3	1.000
4	0.989
5	0.989
6	0.989
7	0.968
8	0.957
9	0.957
10	0.957
11	0.946
12	0.935
13	0.935
14	0.935
15	0.925
16	0.925
17	0.925
18	0.925
19	0.914
20	0.914
21	0.914
22	0.892
23	0.892
24	0.892
25	0.871
26	0.871
27	0.871
28	0.860
29	0.860
30	0.860
31	0.860
32	0.860
33	0.860
34	0.849
35	0.849
36	0.849
37	0.849
38	0.849
39	0.849
40	0.849
41	0.849
42	0.849
43	0.849
44	0.849
45	0.849
46	0.849
47	0.849
48	0.849
49	0.849
50	0.839
51	0.839
52	0.839
53	0.839
54	0.839
55	0.817
56	0.806
57	0.806
58	0.806
59	0.806
60	0.806
61	0.806
62	0.806
63	0.806
64	0.806
65	0.806
66	0.806
67	0.806
68	0.806
69	0.785
70	0.785
71	0.785
72	0.785
73	0.785
74	0.785
75	0.774
76	0.753
77	0.753
78	0.742
79	0.742
80	0.742
81	0.742
82	0.742
83	0.710
84	0.699
85	0.699
86	0.699
87	0.699
88	0.688
89	0.667
90	0.667
91	0.656
92	0.634
93	0.613
94	0.602
95	0.591
96	0.591
97	0.591
98	0.570
99	0.570
100	0.548
101	0.548
102	0.527
103	0.495
104	0.495
105	0.484
106	0.484
107	0.484
108	0.484
109	0.473

FIG. 14. *Duration of apparent tracks.*

DURATION OF REAL TRACKS TO TURN

Minimum number of scans considered = 3

Number of observations = 93

Mean duration = 68.92

Duration more than (scans)	Relative frequency
3	1.000
4	1.000
5	1.000
6	1.000
7	0.957
8	0.946
9	0.935
10	0.914
11	0.882
12	0.860
13	0.828
14	0.806
15	0.806
16	0.806
17	0.785
18	0.763
19	0.742
20	0.710
21	0.688
22	0.688
23	0.667
24	0.656
25	0.656
26	0.645
27	0.634
28	0.613
29	0.613
30	0.602
31	0.602
32	0.602
33	0.591
34	0.581
35	0.581
36	0.570
37	0.559
38	0.548
39	0.548
40	0.548
41	0.548
42	0.538
43	0.538
44	0.527
45	0.516
46	0.516
47	0.516
48	0.516
49	0.516
50	0.516
51	0.495
52	0.495
53	0.484
54	0.484
55	0.484
56	0.473
57	0.473
58	0.462
59	0.452
60	0.441
61	0.441
62	0.441
63	0.441
64	0.441
65	0.441
66	0.441
67	0.441
68	0.430
69	0.419
70	0.409
71	0.398
72	0.398
73	0.398
74	0.398
75	0.398
76	0.387
77	0.387
78	0.376
79	0.376
80	0.376
81	0.355
82	0.355
83	0.344
84	0.333
85	0.333
86	0.333
87	0.333
88	0.333
89	0.333
90	0.323
91	0.323
92	0.312
93	0.301
94	0.301
95	0.290
96	0.290
97	0.290
98	0.280
99	0.280
100	0.280
101	0.280
102	0.269
103	0.258
104	0.247
105	0.247
106	0.247
107	0.237
108	0.226
109	0.226
110	0.226
111	0.226
112	0.226
113	0.226
114	0.204
115	0.204

FIG. 15. *Duration of real tracks to turn.*

DURATION OF REAL TRACKS FROM TURN

Minimum number of scans considered = 3

Number of observations = 35

Mean duration = 97.69

Duration more than (scans)	Relative frequency
3	1.000
4	1.000
5	1.000
6	1.000
7	1.000
8	1.000
9	1.000
10	1.000
11	1.000
12	0.971
13	0.971
14	0.971
15	0.971
16	0.971
17	0.971
18	0.971
19	0.971
20	0.971
21	0.971
22	0.971
23	0.971
24	0.971
25	0.971
26	0.971
27	0.971
28	0.971
29	0.971
30	0.971
31	0.971
32	0.971
33	0.971
34	0.971
35	0.971
36	0.971
37	0.971
38	0.971
39	0.971
40	0.971
41	0.971
42	0.971
43	0.971
44	0.971
45	0.971
46	0.971
47	0.971
48	0.971
49	0.971
50	0.971
51	0.971
52	0.971
53	0.971
54	0.971
55	0.943
56	0.914
57	0.914
58	0.886
59	0.886
60	0.886
61	0.886
62	0.886
63	0.886
64	0.886
65	0.886
66	0.886
67	0.886
68	0.886
69	0.886
70	0.886
71	0.886
72	0.886
73	0.886
74	0.886
75	0.886
76	0.829
77	0.829
78	0.829
79	0.829
80	0.829
81	0.829
82	0.800
83	0.743
84	0.714
85	0.714
86	0.714
87	0.714
88	0.686
89	0.657
90	0.657
91	0.629
92	0.629
93	0.600
94	0.600
95	0.571
96	0.571
97	0.571
98	0.543
99	0.543
100	0.514
101	0.514
102	0.486
103	0.486
104	0.457
105	0.457
106	0.457
107	0.457
108	0.457
109	0.429
110	0.371
111	0.314
112	0.286
113	0.286
114	0.286
115	0.286

FIG. 16. *Duration of real tracks from turn.*

contains detailed information on all Apparent tracks and their break-down into Real tracks. The numbers on the right indicate sequential numbers for the Apparent tracks. Missing numbers belong to Apparent tracks with a duration shorter than the chosen parameter MA. These tracks are not printed out.

Other information on this output shows the slot in which the Apparent tracks were developed, the aircraft number of the associated Real tracks, the duration of Real and Apparent tracks, indicators showing whether Real tracks were initiated and terminated before or after the turn and furthermore, the number of scans in the turn which were left, in case the track is lost in the turn. There is appreciably more information available which is of only minor interest in this context.

The survival-distribution generator tabulates the survival distributions of tracks in the following three classes: Apparent tracks, Real tracks initiated before the turn, and Real tracks not initiated before the turn. Furthermore, the number of samples and the mean values of each class are given. Finally the weighting factor W and the figure of merit F are calculated.

Figures 14, 15, and 16 show examples of the output.

Figure 10 also shows several other programs which are useful for the detailed study of the behaviour of the track logic under investigation. These facilities will not be discussed in detail here.

CHAPTER 30

TRACKING IN AN AIR TRAFFIC CONTROL ENVIRONMENT

B. W. OAKLEY
Royal Radar Establishment, Malvern, England

This paper considers the essence of the tracking problem; why is it necessary to track, and what are the difficulties? It then considers the differences that arise between a co-operative A.T.C. environment and a military environment on which most tracking work has been concentrated in the past. Methods of co-operative tracking are considered, and the implication for conventional primary radar tracking of improvements in radars and tracking logics are reviewed.

1. INTRODUCTION

Tracking in an Air Traffic Control environment does not seem, at first sight, to have much relevance to an audience principally concerned with a Military or Defence environment. But in fact it is very relevant not only because of the safety and control of military aircraft in a civil environment, but also because the detection of the enemy aircraft in an area where there is a predominance of normal air traffic is an acute defensive problem.

2. WHY TRACK?

Before considering how to track aircraft it is essential to consider why tracking is required at all? In the past the answer to this has been so obvious in a military environment that the question has rarely been considered, so that the implications of a shift to the largely co-operative A.T.C. environment may be missed.

Tracking is required for two reasons:

(a) To maintain the current position of aircraft in track stores containing other data on those aircraft. (There are a very few occasions, such as in the case of conflict avoidance, that aircraft position, with no other data, is of use; in most cases the aircraft's position is only of use when associated with other data, such as identity, velocity and intention.) This can be, conversely, considered as maintaining the track identity of each mark on a radar display.

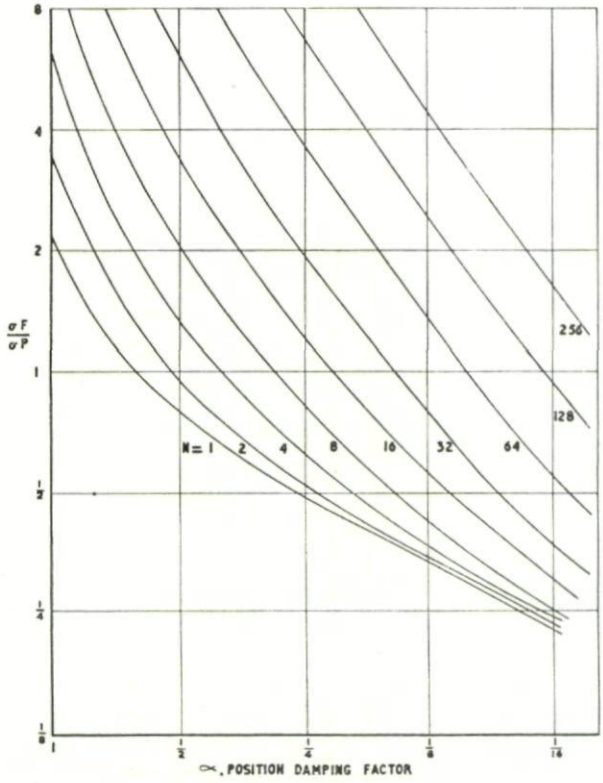
(b) To provide the current velocity of aircraft. For example the velocity of a target aircraft is required for interception calculation purposes.

It should be noted that though the velocity of an aircraft can always be obtained, however imprecisely, from an aircraft track, it is not essential that

this be done. But, in the defence situation it is highly likely that it will be calculated since it will be required in order to carry out the tracking.

With primary radar data only, the plot association can best be carried out by predicting where the aircraft will be on the next radar scan, and then seeking the best correlation of plot-to-forecast position; this prediction process requires a knowledge of track velocity in some form. But if the instantaneous plot position is labelled in such a way that the correlation of plots on successive sweeps into tracks can readily be undertaken, then velocity calculation does not become an essential part of tracking, though the need for velocity for other purposes may be the reason why the track is required.

The accuracy with which the current position is determined must depend upon the reason for which it is required, but frequently this reduces to the



σF: Forecast error, standard deviation.
σP Position measurement, standard deviation of error.
N: Number of data intervals over which the forecast is made.
(The curves are drawn for the classical damping equations, using the normal factors).

FIG. 1. *Forecast position error.*

requirement of resolving one aircraft track from another, so that in a low density environment a lower accuracy can be accepted than in a higher. Errors in the derived velocity are directly proportional to the errors in the position measurement. Errors in a forecast of future position are also directly proportional to the position measurement, but one should note how these errors can become magnified if the prediction is over any significant period of time, unless steps are taken to reduce the position measurement errors by smoothing or track damping (see Fig. 1). Indeed one reason for velocity derivation might be that it enables a more accurate position to be obtained by track smoothing.

It is often the case that relative accuracy between two positions in the same vicinity is of much greater importance than absolute accuracy. It is important to remember this when comparing sources of position information.

3. COMPARISON OF THE A.T.C. AND DEFENCE ENVIRONMENTS

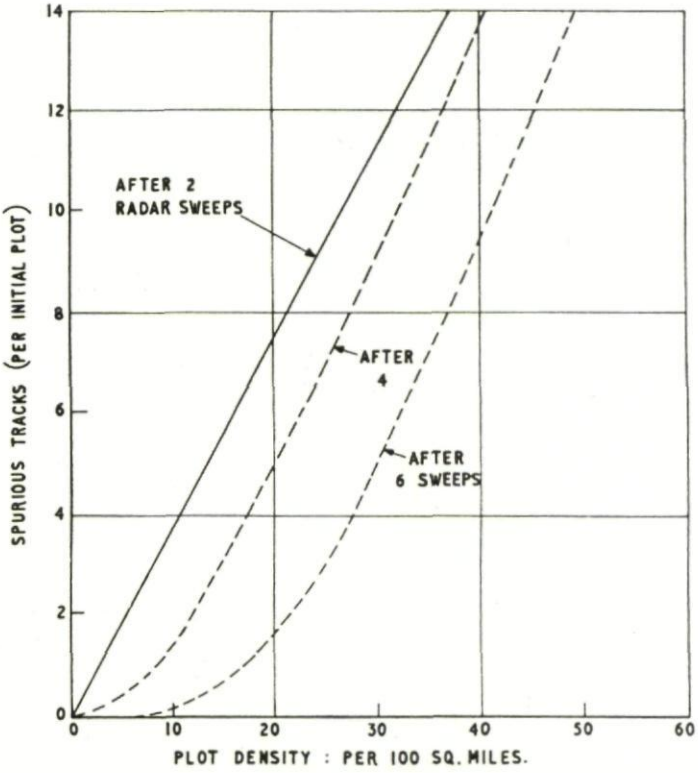
It is a statement of the obvious that the principal distinction between the A.T.C. and the purely defensive environment is that in one all, or most, of the aircraft will be trying to co-operate to ease the tracking task, whilst in the other some will be trying not to co-operate. (It may be noted that in the A.T.C. environment even hostile aircraft are likely to pretend to co-operate since otherwise they might be distinguished by their very lack of co-operation). From the paragraphs above it will be apparent that the essence of co-operation in the tracking task is that the discrete plot positions should carry with them a label by which they can be readily associated. Methods by which this can be carried out are:

(a) By associating a label or code with the radar return. This is the essence of the secondary radar approach.

(b) By generating a message from the aircraft containing its position and identity label. This is the essence of the navigation aid combined with voice report or data link method.

(c) By providing information about the aircraft's flight path in advance that can be used as the basis of the plot association decision. This is the essence of the flight plan method.

These methods will now be examined, in contrast to the defensive primary radar situation where such association data are not usually available. But it should be observed that all these methods have the unfortunate characteristic that a failure of equipment (or man) can result in the loss of data on an aircraft. The normal primary radar method does not suffer from this form of selective failure when operating within its design capabilities. (All methods depending upon electromagnetic radiation suffer from selective failure resulting from loss of cover). Moreover only the primary radar can handle aircraft that are not equipped to co-operate. Evidently a combination of methods is stronger than any one separately.



Notes

1. Drawn for random plot positions and track velocities from rectangular distributions.
2. Maximum track speed = 600 kts.
3. The initial initiation gates are circles of radius set by the maximum track speed expected, and the plot position errors. (3.5 n. mile for 15 second radar sweep.)
4. The subsequent tracking gates reduce in size because of track damping.
5. Tracks fail after one missed plot.

FIG. 2. *Spurious tracks (plan position information only). Position accuracy,* $\sigma P = 0.25$ *n. mile (standard deviation).*

4. CO-OPERATIVE TRACKING TECHNIQUES

4.1. *Secondary Radar Methods*

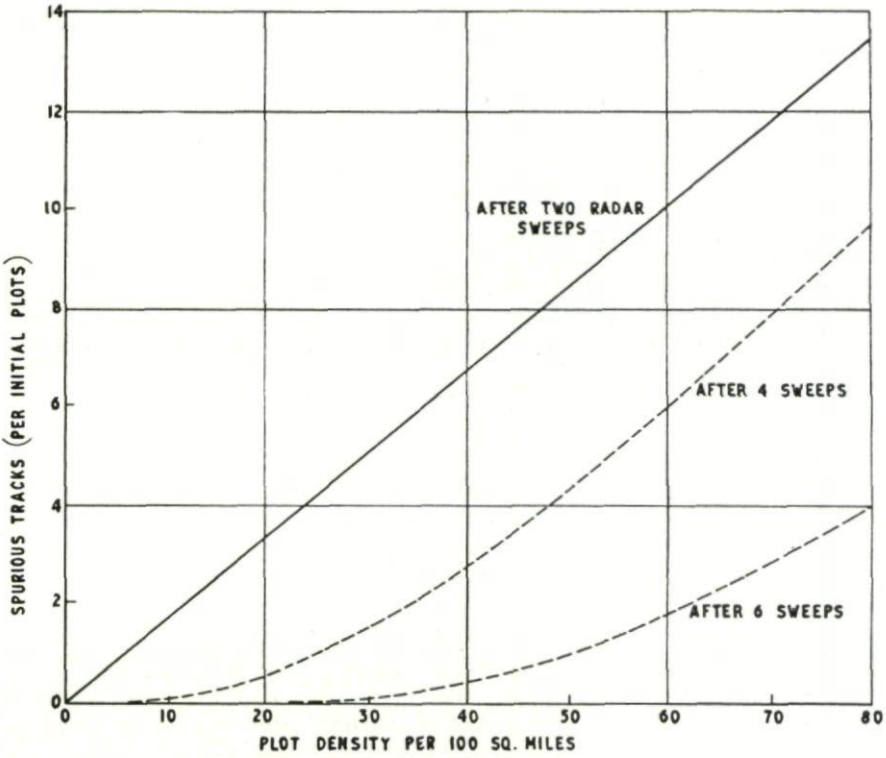
4.1.1 *Clutter Rejection* Secondary radar methods have the important advantage over primary radar methods that returns only come from equipped aircraft, so that problems of ground clutter, meteorological clutter, and angels are not present. Tracking through clutter is difficult because, if there is nothing to distinguish the track from the clutter the association process must inevitably make a virtually random decision, not just on one sweep, but possibly over several successive sweeps; unless there is some extra information, such as flight plan data, to assist in detecting turns, this almost inevitably results in predicting the track through the clutter with a velocity similar to the last correctly observed velocity. A further, and even more serious difficulty with clutter is that it can lead to computer overload by creating a large number of unwanted tracks. This is illustrated in Fig. 2 which shows the way spurious track generation varies with clutter density. Under computer overload conditions even tracks nowhere near clutter may be lost.

4.1.2 *Association by Code* The absence of clutter makes the tracking task much easier; the tracking problem virtually disappears if the returns carry codes on which association can take place. It is not necessary to postulate a complete personnel identity code allocation to make the probability of aircraft with the same code meet in an association difficulty quite low. The probability is inversely proportional to the number of individual codes. Of course care must be taken that the codes really do get randomly distributed.

4.1.3 *Association by Height* In principle the secondary radar returns could be made to carry complete position data derived from airborne navigation equipment. In practise height transponding, using Mode C, is the only form of position data likely to be transmitted, at least for some years to come. If personnel identity is also available then height will not assist the tracking problem (except to the extent that height should really be included as part of the required position data in the tracking task). But if height is available without identity it can prove very useful for tracking in certain circumstances.

The difficulty of tracking on height arises from the non-random distribution of height utilization; aircraft of one type tend to fly at the same altitude if given the choice. Moreover in an area like Europe most of the aircraft are not in level flight, so that a prediction process in height is required if the best use is to be made of the height data.

This tends to mean that the height measurement has to be accurate if it is to be of much use. Figure 3 illustrates the value of height in the primary radar case, (from, for example a stacked beam radar) with standard deviation of height measurement of 2,000′. This is drawn for a particular aircraft height distribution; the situation is not likely to be much improved for any other realistic distribution. In the secondary case height accuracies of 500′



Notes

1 Drawn for same initiation assumptions as Fig. 2.

2 The initial plots are distributed at random with a height density of triangular form, with maximum at 26,000′ and with zero at 0′ and 51,000′.

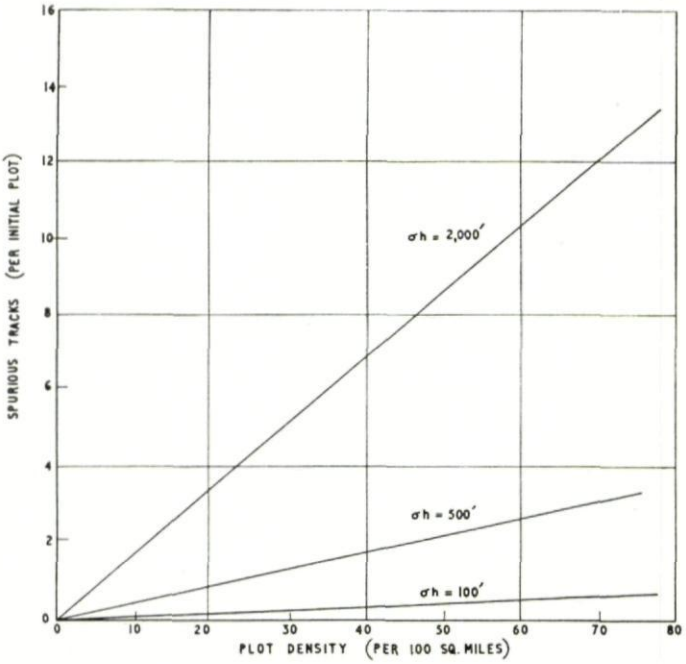
3. Gate in height is taken as $\pm 3 \cdot \sqrt{2} \cdot \sigma h$.

FIG. 3. *Spurious tracks* (*plan position and height*). *Height accuracy* $\sigma h = 2{,}000'$ (*standard deviation*).

(or even 100′) will be available and this will considerably improve the situation (see Fig. 4).

4.1.4 *Effect of Accuracy on Association* The difficulty of association in tracking can be roughly equated to the area of uncertainty over which the association might take place. Now the area of uncertainty resulting from plot position measurement errors will be proportional to those measurement errors. Because of its shorter pulse lengths and wide bandwidths the secondary radar will often be more accurate in range than a primary radar of the present generation. From bandwidth considerations the secondary radar might be expected to have up to a factor of ten times the accuracy, but in practice delays in the transponder system in the aircraft result in accuracies that are only two to three times up on primary radar. Secondary radars tend to have a wider beam width than do primary radars, if only because the primary

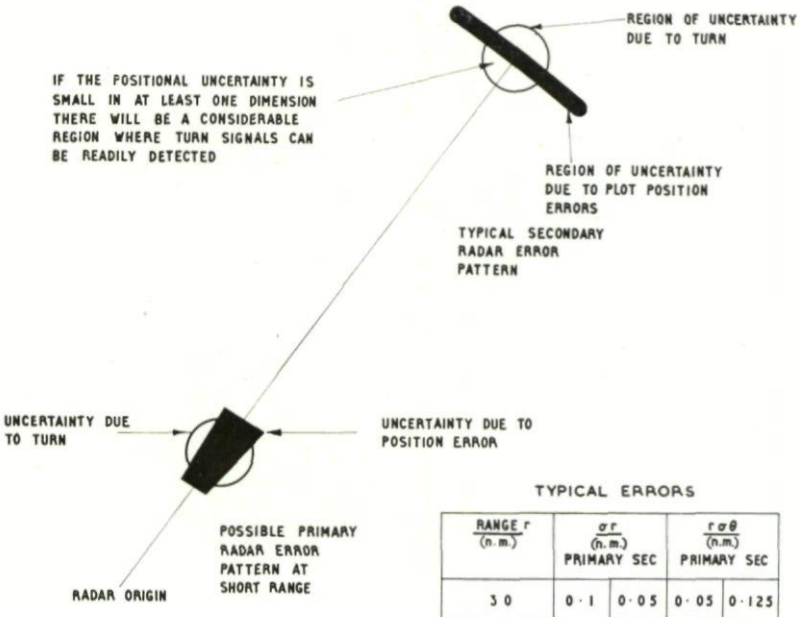
radars are frequently on shorter wavelengths; side lobe suppression on secondary will help here. Effects due to mode interlacing, and aircraft aerial switching also tend to give a lower bearing accuracy to the secondary radar compared with the primary. The effects of range and bearing may roughly cancel out so that the area of uncertainty due to the radar errors is roughly equivalent in the two cases.



Notes
1. Drawn for same assumptions as Fig. 3.
2. For each curve the height gate is taken as $3 \cdot \sqrt{2} \cdot \sigma h$.
3. All curves are for two radar sweeps.

FIG. 4. *Spurious tracks (plan position and height). Illustrating the effect height accuracy,* σh.

But the shape of this area of uncertainty is also important. One of the principal difficulties in automatic tracking is that the magnitude of the "signals" arising from a forecast error due to aircraft manoeuvre is often comparable in magnitude to the "noise" arising from positional errors. Taken over a distribution of track headings relative to the radar origin the surface of turn signal uncertainty will be a circle, whereas the noise is a polar segment (see Fig. 5). Thus there is some advantage in a system where the polar errors in one dimension are small, even if the errors in the other dimension are large, provided the small dimension "noise" is significantly smaller than the turn "signal." The small range errors of the secondary



TYPICAL ERRORS

RANGE r (n.m.)	σr (n.m.) PRIMARY	σr (n.m.) SEC	rσθ (n.m.) PRIMARY	rσθ (n.m.) SEC
30	0·1	0·05	0·05	0·125
60	0·1	0·05	0·1	0·25
120	0·1	0·05	0·2	0·5
180	0·1	0·05	0·3	0·75
240	0·1	0·05	0·4	1·00

FIG. 5. *Effect of polar position errors on turn detection.*

system can be very important at all ranges and bearing, but the small bearing errors of the primary radar only become important at short ranges (usually within the ground clutter region). As an example the secondary range accuracy may be $\frac{1}{20}$ nm compared with turn signals of order $\frac{1}{2}$ nm; primary radar bearing errors (0.1°) will only reduce to $\frac{1}{20}$ nm at 30 nm.

Of course in a particular case of a track flying a radial course this advantage will not be obtained since the turn signal will be into the worst noise direction. But the full advantage will be obtained for the aircraft flying tangentially to the radar.

To take advantage of this, the tracking logic must take account of the polar nature of the measurement errors, and adjust its damping and turn detection thresholds depending upon the tracks position and orientation.

4.1.5 *Problems with Secondary Radar* As with primary radar the cover with secondary radar may not be solid, so that the aircraft can be lost from time to time, even when in a region where good signal returns can be expected. Loss of data in an automatic tracking system is very serious, unless there is some easy method, such as by code, of re-associating the old track with the

new positions after the fade. The difficulty arises because, allowing for aircraft manoeuvre the volume of sky in which an aircraft can be after a fade on a few successive sweeps is usually large compared with the aircraft separations; this leads to the likelihood that the old track will be re-associated with the wrong track. Because of the wavelength the secondary radar ground interference patterns can be serious. The effect of aircraft aerial screening is particularly serious, because the loss of cover tends to occur when the aircraft is in a bank, which is the very time when continuous cover is most needed. It is very difficult to obtain satisfactory cover from only one aerial particularly on military aircraft; two aerials, with peak input signal detection is one solution to this problem, but this leads to an expensive aircraft fit.

Resolution with secondary radar tends to be worse than on primary radar, because the pulse code train extends over 2 miles in range, and the beam widths are greater. In practice the problem is not serious because it is extremely unlikely that two sets of returns will overlap completely in bearing; the undisturbed fringe patterns can be used to obtain the ranges and codes, though with an increased probability of error, and the bearings can be obtained from a knowledge of the overlapped region derived from the code pattern. Moreover there is always a chance, provided the system band widths are adequate of resolving the pulse code trains in range if the returns happen to interlace neatly.

4.2. *Navigation Aid/Data Link Methods*

The combination of a navigation aid and air-to-ground data link is very attractive for co-operative tracking. The airborne navigation aid can often provide accuracies that are considerably better than those derived from ground radar, and of course there is no association problem since the identity can be part of the link message. One other attractive feature is that the data rate can be varied easily, so that data can be provided at a higher rate when, and only when, they are required. (To some extent this can be obtained from electronically scanned ground radars, but the aircraft is more likely to know when the higher data rate is going to be required.) Nor is there any resolution problem, which can be a very important consideration in terminal areas.

The principal difficulty arises from the danger that one aircraft's navigation aid will be in error, relative to that of another aircraft in the same air space. With radar methods relative accuracy is easily obtained since the separate observations are taken with the same equipment. In practice with navigation aids such as Decca it is likely that if a position report from the aircraft can be checked by radar to be within the tolerance of the radar observation then the position report will be correct to a higher accuracy than the radar can determine. Once the data from the navigation aid have been compared with radar data and are being used to form a track, it is possible to measure the apparent jitter on the observations, and this, combined with the detection of any discontinuities, will provide a check on the validity of the data. Evidently a combination of secondary radar with navigation aid/data link techniques would be a very powerful tool.

4.3. *Flight Plan Methods*

It is evident that, just as an aircraft's position is extrapolated between successive radar paints, so it is possible to extrapolate an aircraft's position, with the aid of a pre-notified flight plan, between position reports. But the accuracy of this process is likely to be low, both because of the inherent inaccuracies in the position reports, and because of the long extrapolation time. Any errors in position prediction due to errors in velocity will be proportional to the time over which the prediction has to be made; and since ground speed is the important quantity, wind conditions can lead to considerable position errors at the end of legs. The problem is made more difficult if, as is likely, the aircraft is climbing or descending since its ground speed is then likely to be continually varying.

A further difficulty arises from the doubt about the actual course followed by the aircraft. It may intend to go straight from A to B, but whether it does will depend upon the skill of the crew, the aircraft's navigation equipment, and the weather conditions.

The great advantage of flight plan method over radar methods is that the flight plan contains some knowledge of the future. This can be very helpful to the tracking by helping the association process in times of difficulty. A useful technique is to carry out a running check of the radar data against the flight plan data to see that the track is conforming, within specified limits, with its flight plan. The flight plan can then be used to give warning of heading changes, as a source of velocity when the radar derived velocity is known to be poor (such as immediately after a turn), and as a check on the radar tracking association process.

5. PRIMARY RADAR TRACKING

New techniques, applicable in the Air Traffic Control context, will make the primary radar tracking process much more successful than it has been in the past. These are surveyed here by considering the three major problems in primary radar tracking.

5.1. *Tracking in the Turn*

The problem of loss of data in the turn has already been considered in 4.1.5. In the primary radar case loss is due to the change in echoing area with aircraft aspect. One way of overcoming this is to increase the signal to noise ratio. An expensive way of doing this is to increase the transmitted power or aerial gain. A less expensive way is to improve the receiver noise factor. This is very attractive in the peace time environment where external noise sources are not expected. The addition of a parametric amplifier to a radar receiver may lead to a considerable improvement in performance, for no very great cost. This improvement may be more important in providing more solid cover than in improving the range of cover.

A further improvement in tracking in the turn can be obtained by the use of a variable data rate radar. Higher data rate is not all that much help for

tracks proceeding on largely straight courses, but an improved data rate in a turn would be very helpful. With an electronically scanned aerial it would be possible to step up the data rate on those aircraft which the tracking computer believed to be in a turn, whilst reducing the rate on those aircraft on which there was no association difficulties because of low local traffic densities. One difficulty would lie at the start of the turn, since the tracking system might have difficulty in detecting a turn, and so stepping up the data rate. An extension of the variable data rate radar is the variable dwell time radar where track fading might be overcome by increasing the dwell time on selected aircraft as directed by the tracking computer.

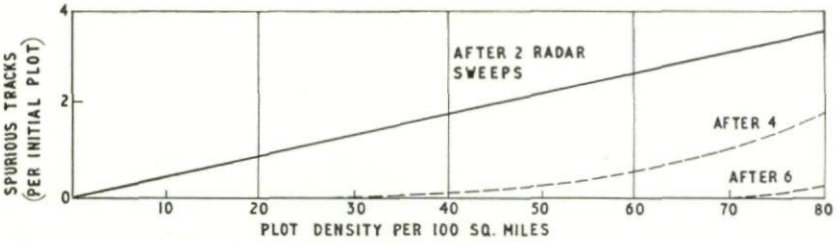
5.2. *Association*

It has been shown that relatively minor improvements in association logic can pay considerable dividends in improving track life. One recent improvement has been to assign probabilities to the likelihood of each possible plot-to-track association, and then to solve the matrix arising from multiple plot and track association possibilities by maximising the overall joint probability. This process can be carried out relatively fast by the application of linear programming techniques.

The process becomes particularly useful in those cases where the probability surface is far from circular. The advent of pulse compression makes possible to obtain extremely accurate range measurement, and this can be a very important factor in improving tracking (see 4.1.4 above).

One fundamental approach to easing the association problem is to reduce the number of occasions on which difficult association decisions have to be made. One way of doing this is to introduce another "dimension" of information that will help to reduce the number of possible plots. In primary radar the extra dimension is likely to be height, or radial velocity. From Fig. 3 it will be seen that radar height measurements are not likely to be accurate enough to make a vast impact on the association problem, unless it happens that the aircraft tracks of interest are well separated from the bulk of the tracks. In practice this may be true in one important respect: about one sixth of the traffic over the U.K. is in the upper air, about one third in the middle air, and one half in the lower air. Even quite crude height finding methods will help to separate the bulk of the traffic in the lower air from that in the upper, and to a lesser extent, in the middle air.

Measurement of radial velocity is an excellent way of overcoming the association problem. Figure 6 shows the improvement to be gained by adding quite crude radial velocity information. Such information might be obtained from a conventional pulse radar without excessive complication by using a two pulse system. Radial velocity is particularly useful because it not only provides a plot filter in another dimension, based on the magnitude of the radial velocity of the plots from sweep to sweep; it also provides a filter on the track association process since its vectorial properties means that it can be used as a further restriction on the association areas (see Fig. 7). This is particularly important in the initiation stages, that is any time when track velocity is not well defined.



Notes
1. Drawn for same initiation assumptions as Fig. 2 except for speed.
2. Speed distribution taken as triangular with maximum at 360 kts. Falling to zero at 176 kts and 544 kts.
3. The tracks are at long range from the radar.

FIG. 6. *Spurious tracks (plan position and radial velocity). Radial velocity accuracy* $\sigma\dot{r} = 12$ kts *(standard deviation).*

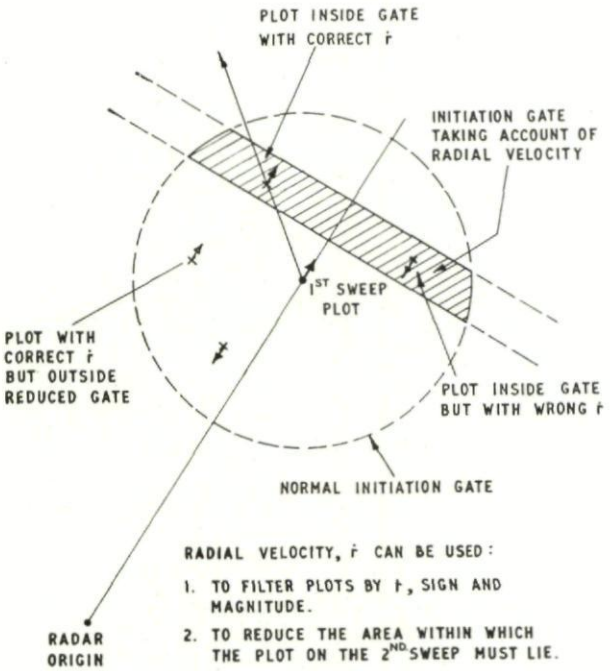


FIG. 7. *Effect of radial velocity on plot association.*

5.3. *Clutter*

As has been discussed in 4.1.1, clutter provides the most important single problem of automatic tracking at short wavelengths. Circular polarization, pulse compression, and pulse length discrimination, and M.T.I. methods help to reduce the effects of clutter; and careful plot extractor design, such as employing area background threshold control, can also help. But clutter remains a problem, particularly if automatic initiation is required as well as tracking.

Improvements in the properties of storage tubes, particularly in the respect of resolution, make sweep to sweep cancellation processes using storage tubes as the storage element attractive. The block diagram for such a system is illustrated in Fig. 8. With this method established tracks will continue to be fed with data, including clutter, but initiation will only take place in areas of low clutter density. The amount of clutter passed to the computer will be controlled by servoing the storage time of the initiation tube from the tracking computer work load.

With the advent of faster computer logic it becomes possible to design computers to control the amount of clutter passed to the tracking computer. One method would be to examine areas of high clutter density, and to measure the distance from a plot to its nearest neighbour; if this separation is less than a threshold set by the tracking computer the plot would be rejected, unless it lay in the tracking gate of an already established track. This

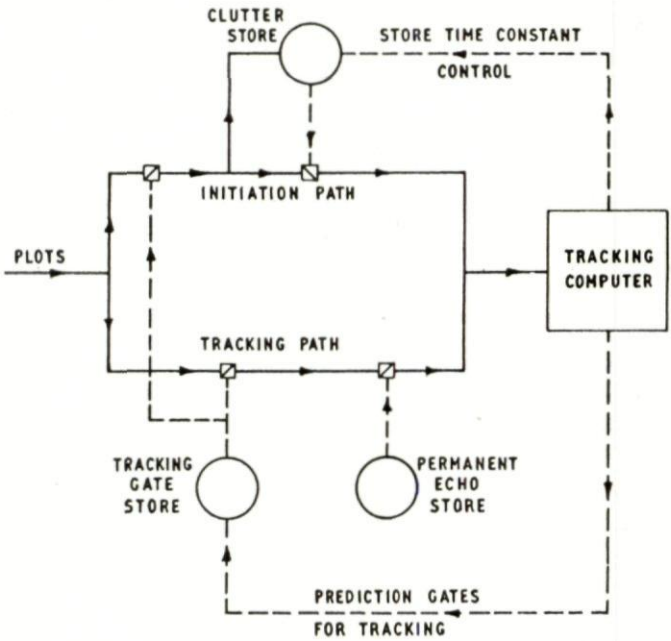


FIG. 8. *Clutter rejection block diagram.*

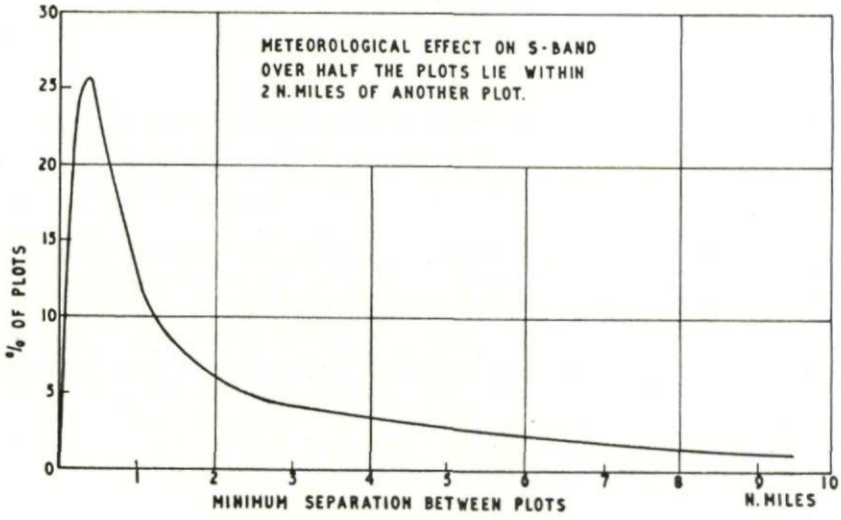


FIG. 9. *Typical clutter separation distribution.*

logic would help to handle the difficult case of clutter that appears in patches of very variable density. Figure 9 shows the variation of plot separation in a typical clutter situation.

6. NATURE OF THE AIR TRAFFIC CONTROL AIR SPACE

The nature of the Air Space Model used in the SHAPE Technical Centre's work described in Chapter 29 is typical of that normally considered for the Air Defence tracking problem. What should the A.T.C. air space model look like? There is no simple answer to this, but Fig. 10 may help. This is a typical 15 minute picture of the A.T.C. air space as seen from a ground radar at Malvern. By analysing enough data of this type it is possible to form some idea of the environment. Figures 11–14 show the distribution of some of the parameters.

7. CONCLUSIONS

(a) In an Air Traffic Control environment methods of tracking that make use of the co-operation of the aircraft can readily be made to largely overcome the tracking problem. Amongst the most attractive are secondary radar tracking using personal identity for association, and airborne navigation aids with data links.

(b) Because of the remaining deficiencies of the co-operative methods considered, and because of the need to follow the tracks for one reason and another are not co-operating, primary radar tracking is still required to confirm and fill in the co-operative methods.

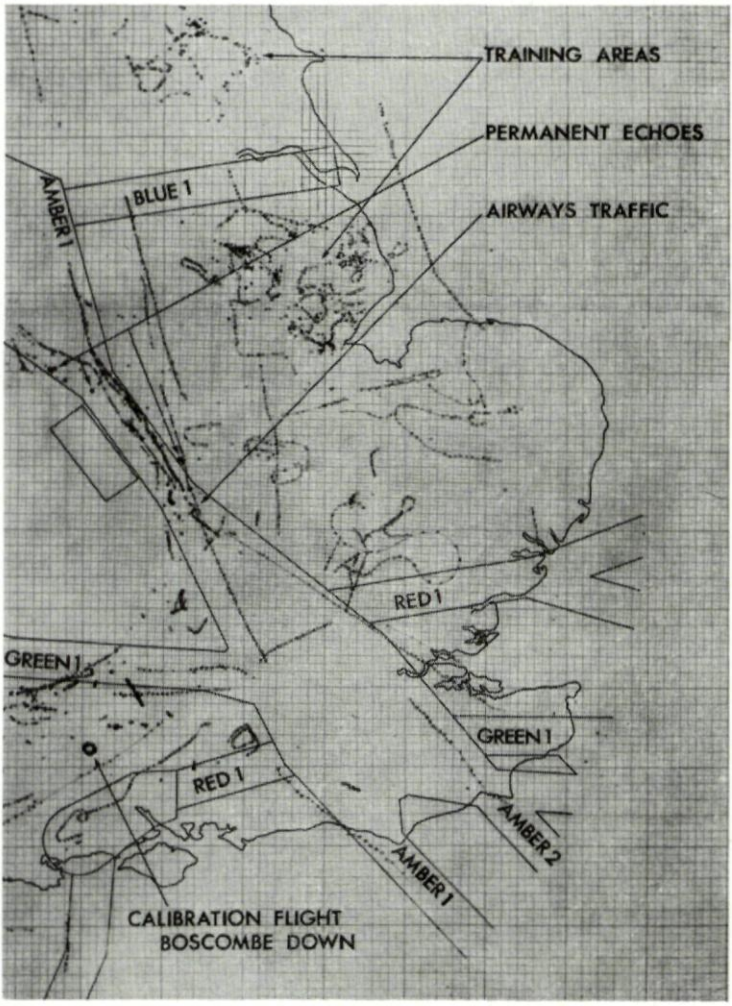


FIG. 10. *Typical* 15 *minute picture of the ATC Air Space as seen from the ground radar at Malvern.*

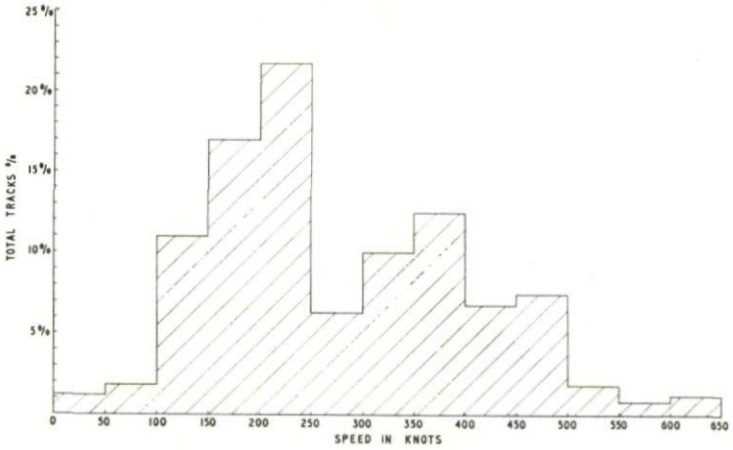


FIG. 11. *Aircraft speed distribution.*

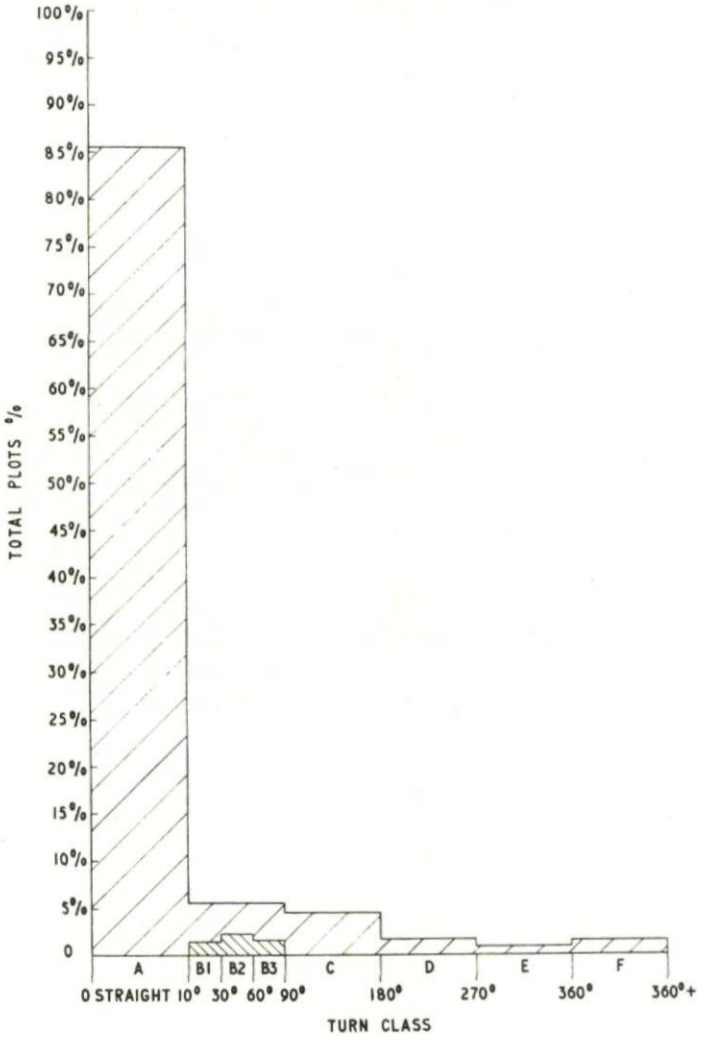


FIG. 12. *Distribution of plots with turn angle.*

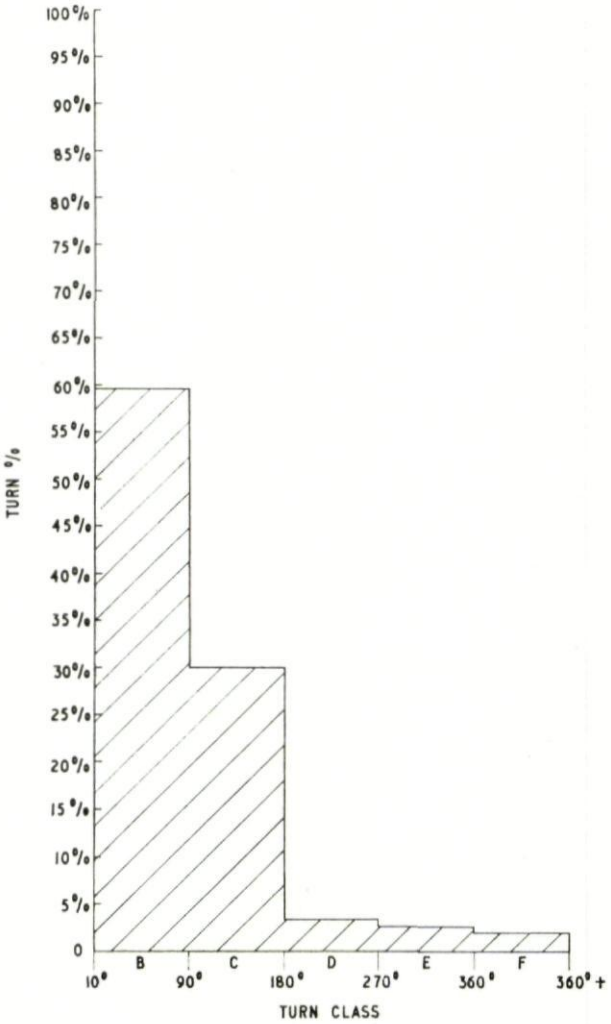


FIG. 13. *Turn angle distribution.*

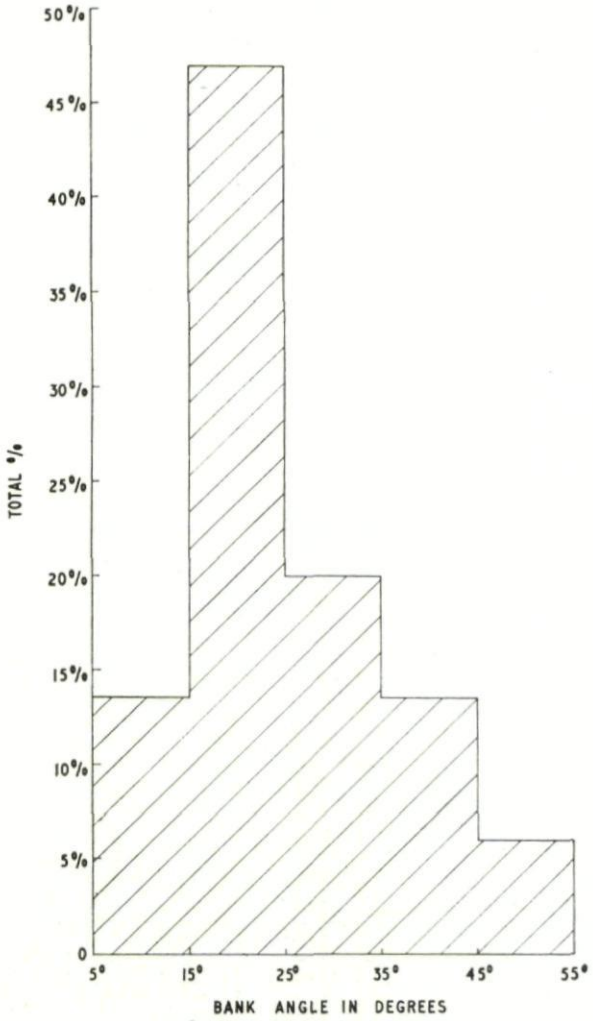


FIG. 14. *Bank angle distribution.*

(c) Recent improvements in radar design have considerably eased the automatic tracking task, but the problems of tracking should be borne in mind in the radar design.

8. ACKNOWLEDGEMENT

The development of the Initiation relationships, used in Fig. 2, 3, 4, and 6, was the work of Mr. A. Brown.

DISCUSSION

E. RECHTIN: We at J.P.L. also have found that the use of what we call mission error ellipsoids (the shape of the errors created by radar tracking, expressed in coordinates critical to the end purpose of the mission) is very important in initial radar design. Such use can show that for certain missions, Doppler is far more valuable than angle and that certain station spacings obviate the need for ranging.

————: You show in the right hand corner of Fig. 5 an area of uncertainty for the secondary radar. This is shown as a fairly wide azimuth tolerance and a narrow range tolerance and on top of that you superimpose a circle which was a manoeuvre uncertainty. This was centred in the azimuth sector.

It would seem to me that if you did not know where the aircraft is within the azimuth sector and you don't know what its going to do, then the total area of uncertainty should extend to the diameter of the circle at one end of the azimuth circle, in other words, you end up with a ray track pattern the diameter of the circle wide.

B. W. OAKLEY: I think your point is quite valid because what one actually sees is the compound effect of those two surfaces put together. But I don't think this invalidates my point that if in fact you get a measurement which falls, in perhaps what I could call the *white* area of the circle, then it is highly likely, though of course never certain, that that is due to a *turn* manoeuvre.

CHAPTER 31

PROVISION AND USE OF ALPHANUMERIC DISPLAY OF PRIMARY AND SECONDARY RADAR TRACKING DATA IN EXPANDING AIR TRAFFIC CONTROL SYSTEM CAPACITY

L. E. SHOEMAKER
Federal Aviation Agency, Washington, D.C.

1. INTRODUCTION

As traffic flow and density increase, the need for air traffic control (ATC) capability improvement through automation of control functions will exist. Not all ATC facilities, however, will experience the same growth or handle the same flows and densities of traffic. Hence, it is necessary to provide several levels of operational capability, each commensurate with the needs of a particular type facility. Thus, automation requirements will vary from those for the busiest center facility downward to those terminal facilities for which primary and secondary radar is provided with little or no data processing beyond conventional plan position data display. Those functions for which automation will be provided include:

(1) primary and secondary radar video data digitizing for remoting via narrow-band width data channels
(2) flight plan aided tracking
(3) generation of alphanumeric data blocks
(4) mosaic display generation
(5) flight plan processing
(6) flight plan distribution and printing
(7) flight plan updating
(8) conflict prediction and resolution
(9) terminal area sequencing
(10) advanced planning flow control.

While the number of primary and secondary radar inputs and the traffic density will vary between centers, automation of all of the above functions is planned in each center generally in the order in which listed. This leads to a single equipment configuration in which only the size of the equipment components varies according to load capacity requirements.

In terminal areas, automation generally will be limited to the generation of alphanumeric data blocks except for some few high density, metroplex terminal facilities that will require the same degree of automation as in the centers.

To satisfy the foregoing automation requirements in the most economically feasible manner, two basic equipment configurations, called the Terminal and Center-Metroplex configurations, have been selected for progressive system implementation of automation in the United States.

The Terminal configuration will have limited functional capability. It will be able only to provide alphanumeric display capability on those aircraft which are visible to the ATC subsystem primary and/or secondary radars. No further expansion of its functional capability will be feasible. The Terminal configuration will be installed at those terminal facilities where the simultaneously airborne controlled traffic is not expected to exceed 100 aircraft and automation functions other than alphanumeric (A/N) display are not justified. Inputs for the Terminal configuration will, in nearly all cases, be derived from one primary and one secondary radar.

The Center-Metroplex configuration, will permit expansion to full functional capability; i.e., when all implementation phases are completed, it will be able to perform all functions as listed above for which automation is presently planned.

This configuration will be installed at all centers and those metroplex terminal facilities where the traffic load of simultaneously airborne controlled aircraft will exceed approximately 100 flights, where more than one primary and secondary radar input is required, or where there is a need for performing additional functions such as flight plan processing, conflict prediction and resolution and terminal area sequencing.

A simplified block diagram of the terminal configuration is shown in Fig. 1. Inputs will be derived from primary and/or secondary radars. These input data will be fed to the controller's display through independent video and digital channels. In the video channel, the data will be passed through scan converters and applied to bright, flicker-free, television type displays. In the digital channel, video will go first to a data digitizer. In the digitizer the center of the video return will be determined and digital signals representing this position generated. Digitized primary radar position data are fed to the computer for tracking. After positive radar identification is established, tracking will be initiated on controlled non-transponder equipped aircraft by entering the track instruction and the associated aircraft identity and cleared altitude into the computer through the controller's keyboard. Once the target position, approximate heading and velocity are entered into the computer a track will be established from which a prediction can be made of the most probable target position on successive radar scans. The computer will use this prediction to continue the association of identity and altitude data with the aircraft target.

Two inputs are taken from the SSR data portion of the digitizer; one to the tracking computer and one to the display processor. Thus, digitized position data on all transponder equipped aircraft will be available to the computer. The only fraction of these data stored by the tracking computer, however, will be related to those aircraft on which the controller initiates tracking. Normally, this will be only on those aircraft carrying a 64 code transponder providing a non-discrete reply and no pressure altitude data.

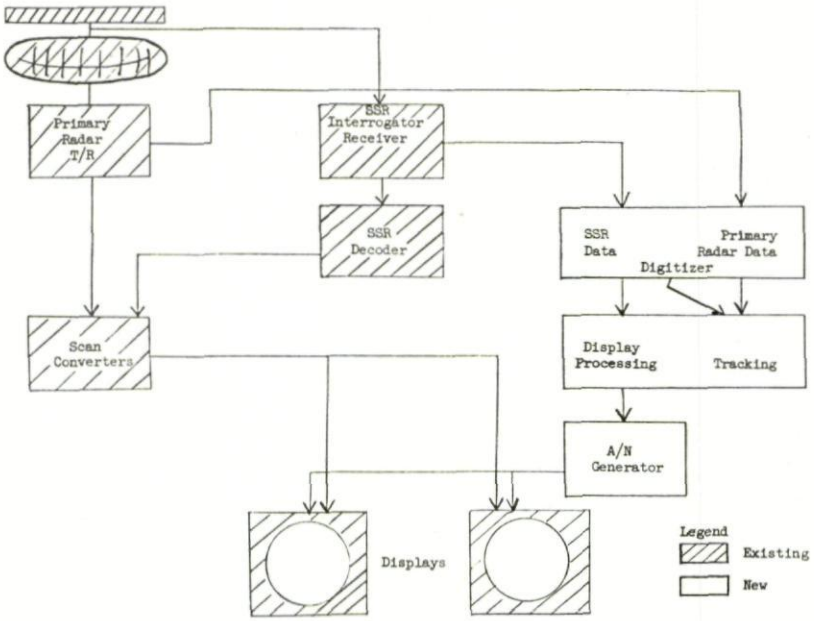


FIG. 1. *Terminal equipment configuration.*

The initiation of tracking and insertion into the computer of identity and altitude or flight level data will be accomplished through the controller's keyboard. For transponder equipped aircraft on discrete codes, the digital position signals and associated pressure altitude and identity codes will be transferred to a display processor. The processor will search the display store, locate the proper aircraft identity if it is in the store, format the aircraft identity for display, and transponder derived altitude or flight level data available on the entered alphanumeric data. If there are no stored data available on the aircraft, the processor will format the identity code and altitude or flight level for display. If an aircraft is operating at an assigned or hemispherical altitude the transmitted standard pressure altitude data will be converted to actual altitude before display. If "flight level" is being flown, barometric pressure compensation will be unnecessary since the transmitted pressure altitude data is already in the proper form. The alphanumeric display formats will then be fed to the A/N generator and, from there, to the display.

A keyboard (not shown in Fig. 1) will be provided at each console to permit the controller to:

(a) select for display all aircraft between certain altitudes;
(b) select for display those aircraft replying to selected non-discrete identity codes;
(c) select for display by aircraft identity only those aircraft with discrete codes, as desired;

(d) insert or delete climb or descent arrow, cleared altitude/flight level information or handoff markers (for inter-sector handoffs).

It may also be used to:

(e) insert aircraft identities and associate them with discrete identity codes when such data have not been previously entered.

In the lower activity terminals, wired-program data processing machines and conventional plan position displays may be used to provide comparable functional capability but with somewhat less flexibility.

Before discussing the implementation of the Center-Metroplex configuration, it is necessary to describe the computer which this uses.

2. COMPUTER FOR THE CENTER-METROPLEX EQUIPMENT CONFIGURATION

A simplified block diagram of a computer complex and its associated peripheral equipments, suitable for field implementation as part of the Center-Metroplex equipment configuration, is shown in Fig. 2. Depending upon site location, the required number of primary and secondary radar inputs will vary along with computer capacity requirements. The computer complex is sufficiently flexible that it can be configured to handle a range of traffic loading up to 450 simultaneously airborne controlled aircraft. The total computing load will, of course, be determined by the nature and number

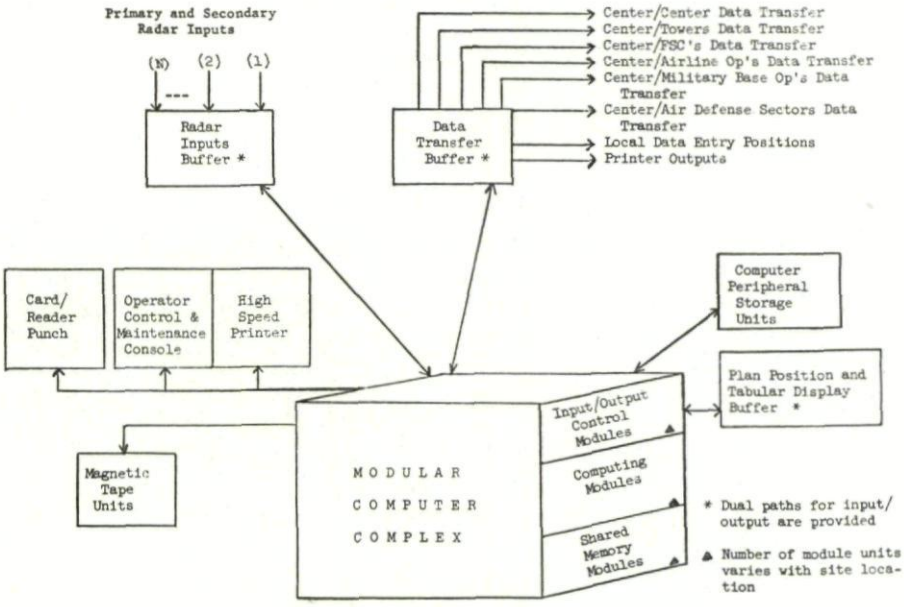


FIG. 2. *Typical computer system configuration (simplified).*

of ATC functions to be performed automatically. In keeping with our progressive field implementation plan, the computer configuration will have the capability not only to grow with traffic increase, but to permit automation of additional ATC functions at a minimum increase in cost.

Two factors are of prime significance when considering the type of data processor needed to satisfy the requirements of the various types of ATC facilities to be equipped with the Center-Metroplex configuration. These factors are:

(a) the multiplicity of jobs to be done by the data processor

(b) the wide difference between the two extremes of facility size

Examination of the first factor indicates that a data processor of very large capacity and/or high speed will be required. Since circumstances prevent implementation of a final phase of progressive improvement for several years, it is desirable to start with a processor having considerably less capacity than that ultimately required but that will permit the addition of hardware and software at a later date.

Accepting as a premise that standard software packages are necessary to the efficient implementation of the Center-Metroplex computer, it follows that standardized hardware will also be necessary. Looking at the wide difference in facility sizes, it becomes obvious that, in all phases, data processors of widely varying size will be needed. Thus, a processor which can be installed following the modular building block principle is a logical choice. Also, to simplify logistics support (spare parts, personnel training and software implementation), it is desirable that the same basic modules be used in all facilities and that the number of modules vary with facility load requirements and phases of implementation.

3. COMPUTER CHARACTERISTICS

Having concluded that a modular building block type computer is most suitable, it then became necessary to select other characteristics needed to meet specific operational requirements, such as reliability, versatility and ease of maintenance. Some of the specific characteristics selected are:

(a) Step-by-step increase in processor capacity will be accomplished by modular expansion in the three major areas of computing speed, active shared memory capacity and input/output transfer capability.

(b) No single modular unit will be critical to system operation. The design will, therefore, ensure that failures affect only individual modules and, while system performance may be degraded, a total processor failure will be highly unlikely. For example, a single input/output module failure may degrade the transfer rate capability of the system, but will not affect a peripheral device, computing speed, or active shared memory capacity.

(c) The system will have active shared memory modules to which any computing module has access. "Shared memory" modules are high

speed, random access units available to every on-line computing module within the normal memory access cycle time.

(d) Program routines and associated data necessary to the performance of operational functions will be controlled by a single executive program distributed in shared memory. The executive program will assign tasks to individual computing modules according to the first available, hence no single computing module will be critical to the system operation.

(e) The processor will monitor the status of its major components and define its capability by reference to a pre-established matrix function table. (On-line elements versus functional capability). This will facilitate rapid recovery from failure and permit balancing of the functional capability with an order of priority according to available on-line components.

(f) The processor will have excess capacity on a modular unit basis and be able to utilize this capacity under program control to provide fail-safe and fail-soft modes of operation. This concept is preferable to that of having active and switchable standby units in that all modules function on line, the resulting excess capacity being available in a manner that permits fast recovery from an individual module failure.

(g) The modular configuration will allow use of excess capacity, beyond that required to maintain reliability during periods of light traffic activity to perform off-line non-operational functions without interference to the on-line operational functions. Examples of off-line non-operational functions are program check out, statistical analysis and data reduction. The excess capacity, however, should be immediately available to the operational program.

Fail-safe and fail-soft modes of operation permit uninterrupted operation when successive module failure occurs, albeit at reduced capacity. "Fail-safe" means sufficient capacity that all functions will continue to be performed at the rate required, even though some modules are inoperative. "Fail-soft," on the other hand, refers to that condition when failures have occurred to the extent that fail-safe operation is no longer possible. In this case, one or more modes of fail-soft operation are possible. A level may be reached where all functions are still being performed, but certain of the less critical ones are being performed at a slower rate.

A lower level may be reached where the less critical functions are no longer being performed. At an even lower level, higher priority functions will be performed less often, and so on.

4. CENTER-METROPLEX CONFIGURATION

The equipments used for the initial phase of implementation of the Center-Metroplex configuration are shown in the simplified block diagram of Fig. 3.

The number of primary and secondary radar inputs is not limited as is the

case with the Terminal configuration. The number will be determined by the amount of coverage required for a given facility, and the buffering system will accommodate any needed variations in the number of input/output devices.

The primary and secondary radar video inputs will be converted to digital form at the radar sites for transmission to the control facilities. The modular computer capacity will accommodate flight plan aided radar tracking, automatic flight plan updating of radar targets representing controlled

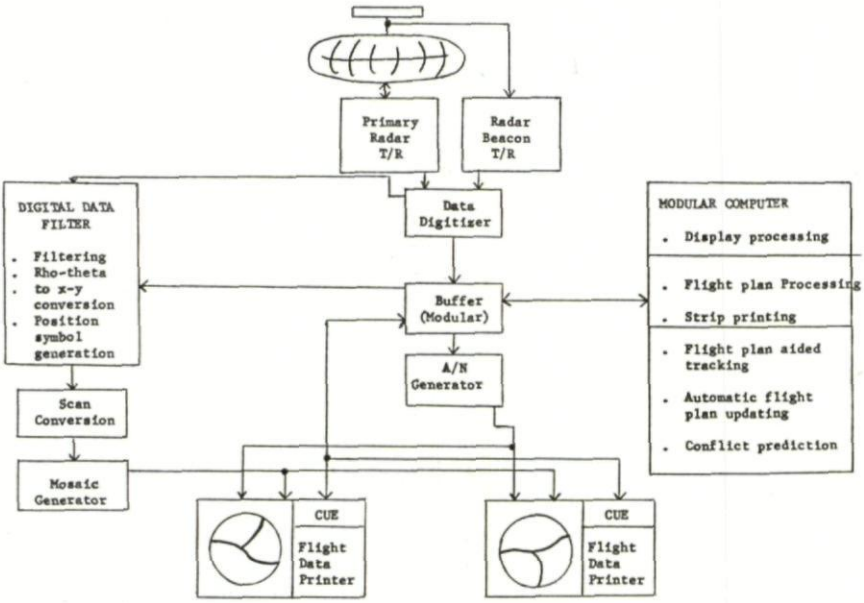


FIG. 3. *Phase 1 implementation of the Center-Metroplex equipment configuration.*

flights, and conflict prediction functions. In addition, a mosaic plan view display capability and a digital video channel will be incorporated into the system. Flight data printers and controller operated computer updating equipment (CUE) will be installed in the control consoles.

The mosaic plan view display is obtained from several primary and/or secondary radar inputs covering an area. The data are combined in such a manner that the returns from not more than one primary and associated secondary radar are displayed in a particular area of the display and, further, the returns from two different radars are not portrayed in any area of the display unless that area represents a boundary between two or more separate areas. The mosaic technique utilizes a common TV video time base related to the common generation, and is done on the readout side of the scan converters being accomplished by inserting the proper time delays at the start of each TV line and frame sweep.

A common time reference is required for all plan view displays so that the

computer is able to convert aircraft position from the common coordinate system used in the computer to individual display coordinates. A common time base results in one large mosaic picture properly related to the computer generated alphanumeric data.

In the video channel, a digital data filter is used. This unit accepts the combined primary and secondary radar digital message output data transmitted from the radar site and selects data for display in accordance with controller activated switch settings. It also changes the position coordinates of selected targets from rho-theta to *x-y* form and generates appropriate position symbols. In the data digitizer at the site coincident primary radar returns are suppressed on those aircraft from which valid SSR responses are received. Accordingly, the digital data filter will pass only those SSR targets selected for display and those primary radar targets not suppressed in the digitizer. However, the SSR target return message from the digitizer will indicate when a coincident primary radar target return existed so that the controller may display all primary radar targets if desired. The digital data filter will accept radar target or flight plan position information from the computer on aircraft which were selected either as a result of controller request or the execution of a computer program. This channel is also utilized as the fail-soft channel from the computer to the display when the A/N generator becomes inoperative.

The flight data printer will print flight progress strips for each flight as appropriate to the circumstances. For instance, in the case of a departure from within a center's area, strips will be printed for the departure control position. Upon departure of the flight, its takeoff time will be entered into the computer via the CUE keyboard. Immediately, the computer will calculate estimates over subsequent fixes along the flight's planned route, and cause strips to be printed at all sectors concerned. Simultaneously, the computer can direct a current flight plan message to the next center en route. As the flight progresses, subsequent changes in route or altitude assignments will be inserted through the CUE keyboard. The resulting updated information will be presented on CUE indicators in all sectors affected by the changes.

The computer program will also include provisions for conflict prediction and fail-soft system operation. Short term conflict alert will be presented on the plan view display. A conflict alert symbol will be associated with the respective aircraft targets, and details will appear in an area of the display not used for traffic display. Long term conflict alerts will be presented similarly, except that the detailed description may be printed out on the flight data printer, with only the alert symbol appearing on the plan view display.

Four modes of system operation, shown in the chart of Fig. 4, will be available with the Center-Metroplex equipment configuration.

In positive control sectors, where all traffic is controlled, a fully synthetic display derived from the computer A/N channel represents the full capability, normal mode of operation. Computer generated windows will allow display of primary radar data on controlled aircraft whose transponders

have failed. The video channel provides a back-up capability in the event of component failure in the computer A/N channel to permit uninterrupted traffic control with radar separation.

In non-positive control sectors providing radar advisory service, both the computer A/N and video channels are utilized during full capability normal operation. Alphanumeric data are displayed on controlled aircraft and video channel position symbols are displayed for both controlled and uncontrolled aircraft. Additionally, position and altitude data will appear on

Operating condition	Type of sector	Mode and capability	Channel routing of computer window data	Data displayed
Normal	Positive control Non-positive control	Normal/ Full Normal/ Full	A/N	Computer position symbols and alpha-numerics Video channel position symbols and computer channel position symbols and alpha-numerics
A/N Gen. inoperative	Positive and non-positive control	Fail soft/ reduced	Video	Video channel position symbols including some generated by the computer
Computer inoperative	Positive and non-positive control	Fail soft/ reduced		Video channel position symbols only
Video channel inoperative	Positive control Non-positive control	Normal/No video backup Fail soft/ reduced	A/N	Computer position symbols and alpha-numerics Computer position symbols and alpha-numerics

FIG. 4. *Table showing four modes of system operation for the Center-Metroplex equipment configuration.*

those uncontrolled transponder equipped aircraft transmitting pressure altitude which are within the airspace under jurisdiction of these sectors. Computer generated flight plan positions will be displayed either in response to controller request or as the result of execution of a computer program.

Three fail-soft modes, providing reduced system capability, will be available in both positive or non-positive control sectors. One fail-soft mode will be used in the event of A/N generator failure. In this mode, video channel and selected computer generated position symbols are displayed. The computer generated data are routed through a fail-soft path to the video channel and are displayed using the video channel symbol generator. A second fail-soft mode will permit the display of controller selected video channel position symbols only, should complete computer failure occur.

A third fail-soft mode will be available should a failure occur in the video channel. Since the total display in positive control sectors is normally

synthetic, operation will be unaffected but there will be no video back-up capability. In non-positive control sectors the quality of radar advisory service may either be reduced or this service may be terminated if the capacity of the computer A/N channel is exceeded due to heavy traffic loading. In this mode the computer will be called upon either to present all target data via the A/N channel to displays suffering the loss of video input, or to evaluate the potential threat each offers to controlled aircraft and present only those on which radar advisories should be given. Eventually the computer A/N channel capability should be expanded to handle the full requirements of radar advisory service without reliance on the video channel.

5. CENTER-METROPLEX CONFIGURATION-IMPLEMENTATION, PHASE 2

As a later step of the Center-Metroplex configuration, a tabular display system will replace the paper strip flight progress boards and CUE displays. A simplified block diagram is shown in Fig. 5. The display system will interface with the computer through the display buffering system. It will consist of central buffering and character generation equipment, tabular displays and keyboards located at both the active control and new planning consoles. The latter will include multi-sector traffic flow planning consoles, arrival/departure planning consoles for terminal traffic and combined supervisory and flow control console.

Evaluation of the hardware and software required for the tabular display system will be conducted at our National Aviation Facilities Experimental Center at Atlantic City, New Jersey. Computer programs will include the functions of advanced planning, terminal area traffic sequencing, flow control and valid solutions to potential conflict situations.

The data displayed will be determined by the system mode of operation, as shown in the chart of Fig. 6. During normal operation, abbreviated flight data on all controlled aircraft and the necessary clearance data, including rerouting, will be presented automatically on the electronic tabular display located at each active controller's console. The plan view display will continue to allow the display of graphic route information on selected controlled aircraft and will automatically display sequence orders and conflict alerts. Computer recommended solutions to potential conflicts may be called up on the plan view display. Computer generated sequence orders, conflict alerts and solutions will be presented on the tabular display automatically. Flight data printers will be available for the controller to request complete flight plan, conflict solution or clearance data on selected aircraft. Also, without controller request, the printers will periodically print abbreviated flight plan data on all controlled aircraft to provide a record in the event of complete computer failure.

During the fail-soft mode of operation when the plan view display A/N generator is inoperative, computer generated sequence orders and conflict

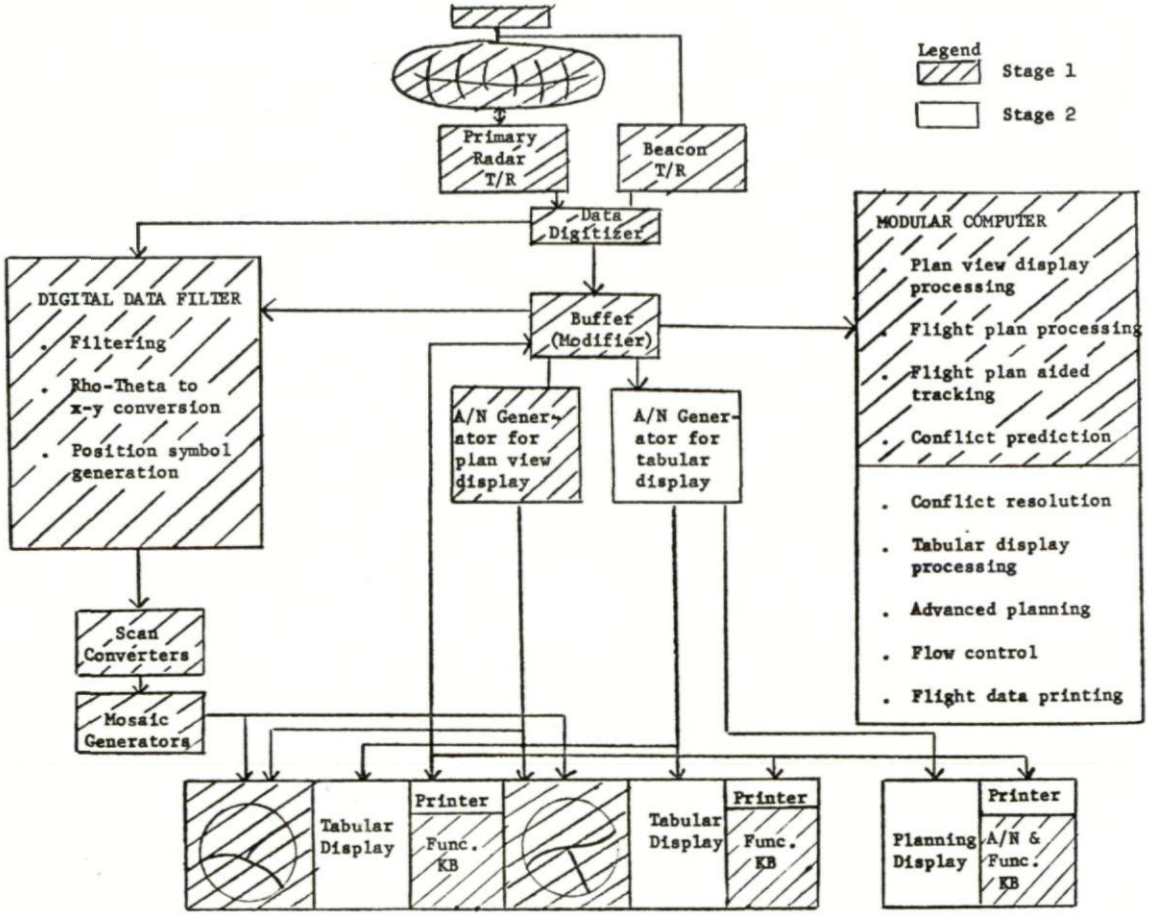


FIG. 5. *Phase 2 implementation of the Centre-Metroplex equipment configuration.*

Condition of operation	Electronic tabular display	Plan view display	Flight data printer
Normal system operation	.Abbreviated flight plans on all controlled aircraft .Clearance data (including rerouting) .Sequence orders .Conflict alerts and solutions .Needed symbology	.Aircraft position symbols .A/N data .Graphic route segments on selected aircraft .Conflict alerts and solutions .Sequence orders	.Complete flight plans on selected aircraft .Abbreviated flight plan data on all controlled aircraft—renewed periodically .Complete conflict situation on selected aircraft .Complete clearance data on selected aircraft
Plan view display A/N generator inoperative	.Normal system operation	.Video channel data including some computer generated symbols	.Normal system operation
Tabular display A/N generator inoperative	.Data current when A/N generator failed	.Normal operation	.Normal system operation
Computer inoperative	.Data displayed at time of computer failure	.Video channel data only	.Data current when computer failed
Video channel failure	.Normal system operation	.Normal system operation plus data on uncontrolled aircraft until computer becomes overloaded	.Normal system operation
Printer inoperative	.Normal system operation	.Normal system operation	.No new data

FIG. 6. *Display operation during various modes of operation in the Center-Metroplex equipment configuration.*

alerts normally shown on the plan view display will be presented on the tabular display only. No graphic route display will be possible.

During the fail-soft mode of operation when the computer is inoperative, use is made of the abbreviated flight plan data printed by the flight data printer with updating and posting accomplished manually. Additionally, the latest data presented on the tabular display at the time of computer failure will continue to appear, and the video channel will continue to provide symbolic primary and secondary radar data.

Should the video channel fail, the only effect on the displays will be the possible loss of position data on uncontrolled aircraft should heavy traffic

loading cause computer capacity to be exceeded. In the event of failure of the tabular display or its A/N generator, flight plan and clearance data, periodically updated by the computer, will be provided by the flight data printer. Should the flight data printer fail, the tabular and plan view displays will continue to operate normally but without printed copy backup.

6. SUMMARY OF OPERATIONAL IMPROVEMENTS

Discrete flight identity and altitude/flight level data will be provided on the controller's plan-view display. This will be done by automatically associating an alphanumeric tag with selected targets. The flight plan message, forwarded from an adjacent facility, will give advance notice on a particular transponder equipped aircraft expected to enter a facility's area. Aircraft identities and discrete code assignments will be inserted into the computer by means of a keyboard during the initial entry of flights into the system. In effect, for 4096 code transponder equipped aircraft, these entries will tell the computer, for example, that flight KLM 404 is to be associated with code 3472. Thus, once a flight enters a sector, there will be continued automatic association between the SSR target return and the alphanumeric identity and altitude tag displayed as the aircraft transits the sector. Where adjacent facilities are equipped with an automatic system capability, the aircraft identities and associated discrete SSR codes will be transferred automatically between computers. This will enable automatic transfer of radar target identification, eliminating manual radar handoff on discrete code transponder equipped aircraft. Those aircraft without discrete codes may be displayed either symbolically, by code number, or by area discrimination in conjunction with a code employment plan. For aircraft not equipped with transponders, the alphanumeric data will be associated with the aircraft by means of computer tracking once positive target identification and track initiation is performed by the controller.

Implementation of the Center-Metroplex configuration introduces a level of capability which will permit the data processor to check the system and call situations requiring control action to the controller's attention, such as conflict detections. It provides the controller with a "cleaned up" display and a fail-soft system capability that will permit operation in various modes in the event of component or subsystem failure. In addition, it may be possible to introduce assistance in the terminal process in the form of limited final approach spacing incorporating time-to-turn readout to the spacing controller.

An evaluation process will be included in the computer program to permit the data processing subsystem to detect a situation of routine or potentially dangerous nature which requires control action and, through the display subsystem, post a forced alert. Representative situations include transfer of control and conflict prediction.

Conflict prediction is made possible by a marriage of automatic tracking and flight plan processing functions. This results in the introduction of flight plan aided tracking and automatic updating of flight progress. The

latter process provides the basic computer inputs needed for the process of conflict detection; i.e., radar tracker derived velocity and three dimensional position data provide the computer with accurate current progress on controlled flights from which it can extrapolate future positions and detect potential conflicts.

Two forms of conflict prediction will be performed: Short term, which will predict potential conflicts calculated to occur within the ensuing five minutes based on position and velocity and long term prediction, on a coarse basis, projected forward in time by a maximum of thirty minutes based on updated flight plan position and velocity information. Long term conflict prediction will assist the controller in foreseeing concentrations of conflicts which need to be resolved in advance by rerouting or changes in cruising levels. Short term conflict prediction will alert the controller to situations requiring immediate action for resolution usually with radar vectoring.

The capability of extrapolating flight plan positions will also allow presentation on the plan view display of flight plan positions on selected controlled aircraft together with flight plan routes, as needed. By comparison with actual position data, the controller, assisted by the computer, will detect deviations from the flight plan route. In airspace where primary and/or secondary radar coverage does not exist, the controller's display will present extrapolated flight plan positions and tracks which can be updated periodically by manual position entries based on pilot position reports.

The mosaic display capability, using data inputs from multiple primary and/or secondary radar will make available to each controller a display of the area under his jurisdiction independent of the particular radar providing coverage of that area. The mosaic display will be made up of a number of "tiles" with each tile presenting the best coverage of a single primary and/or secondary radar. Using this type display, multiple targets will not appear on individual aircraft except occasionally while aircraft are transiting a boundary between tiles. Suitable controls will permit selection of the desired portion of the mosaic for the most advantageous display of a particular area. Expansion controls, which are part of the display, will allow a double scale magnification of any part of the mosaic. With incorporation of the mosaic capability, the computer program is modified to enable display of A/N data from the same radars superimposed on the video mosaic. The controller is then able to select data for display either from the video channel, the A/N channel or both, irrespective of radar site limitations.

Cleaning up of the controller's display will be accomplished in a number of ways. As a result of installing a digital video channel, small, uniform-size symbols will appear on the plan view display in place of the present-day video returns which are proportionate in size to antenna beam-width. In addition to "pinpointing" the position of each return, these symbols will also indicate whether the return represents a primary radar target, a non-discrete code transponder equipped aircraft, or a selected discrete code transponder equipped aircraft. In the case of SSR targets, the normal symbol will automatically be replaced by other special symbols when, for example, the pilot selects transponder code 7700 to indicate an emergency, 7600 to indicate radio failure,

or actuates the special identification feature. Other symbols will be used to indicate computer information. Background clutter, including that due to weather, will be reduced by a masking process in the digitizer. The outline of the masked area will be shown through the use of clutter map symbols. Further cleaning up will be accomplished by providing a filter control by which the controller will be able to select SSR targets by altitude stratum for display on the basis of transponder transmitted pressure altitude data.

In non-positive control airspace, tracker acquisition of non-discrete secondary and primary radar targets will be facilitated by display of the computer predicted flight plan positions. For the acquisition of targets in positive control airspace an aircraft not responding with a discrete code will be expedited by having a computer generated window automatically appear at a point on the display where the controlled aircraft is expected to enter the sector. The window will be positioned according to the flight plan extrapolated position of the aircraft. This will allow non-discrete secondary and primary radar data to appear enabling the controller to identify the target and initiate tracking. In cases where a tracker becomes disassociated with its aircraft target, reacquisition will be facilitated by permitting the window to reappear in the area where the desired target is predicted.

Provision will be made for the presentation on the controller's plan view display of other keyboard selectable data such as a video map. The automatic display of hazardous weather outlines on the controller's plan view display will be provided. The type of outline displayed will signify the nature and intensity of the possible hazard; i.e., turbulence, ice, etc.

Phase 2 of the Center-Metroplex configuration implementation will be largely related to providing an improved coordination and planning display subsystem to improve efficiency. It will be achieved through added sophistication of the computer program and provision of improved input/output devices. It is important to note that the role of the computer will change, as a result, from one of clerical assistance and situation monitoring to one of providing active assistance in the decision-making process.

Combination tabular display and computer input devices will be human engineered to vastly simplify the problems of communication between the controller and the data processor. These devices will display essential flight plan data in abbreviated tabular form. Integrated with the tabular display will be pushbuttons which light in various ways for discrete computer data input and output. These devices, together with expanded function keyboards, will greatly reduce inter-controller voice coordination. (A function keyboard, as distinguished from an A/N keyboard, will allow input of discrete "canned" messages, such as climb, descend, hold, and reroute.) Simple push-button actions will enable communication to the computer, other electronic subsystems and other controllers simultaneously. Eventually, the same means can be used to convey "canned" ATC messages to pilots in the event of a future ATC data link being used.

Prior to this period, the sector controller will control traffic both on a short term basis to resolve imminent conflicts, and on a longer term basis to minimize potential concentrations of conflicts. These later steps will divide this

workload among "active controllers" for short term conflict resolution and "planning controllers" for long term conflict resolution. While active controllers will be responsible for individual sectors, the planning controller will coordinate the activities of a number of sectors. Accordingly, when a planning controller perceives an unacceptable concentration of conflict situations or abnormal traffic build-up developing in one sector, he will be able to assist active controllers in the rerouting of traffic within the sector or through other sectors or to take other measures to limit the workload in the sector involved.

The computer will be programmed to assist the controller in the decision-making process to the extent that it will recommend solutions to the problems it has detected. These solutions will be in the form of specific control actions that could be taken as valid solutions to the problems. The computer will, for example:

(a) Recommend route and/or cruising level changes to assist the planning controller in reducing the incidence rate of future conflict situations. This will permit expansion of sector size by increasing the number of aircraft that an active controller can safely handle;

(b) Sequence and resequence aircraft entering metroplex terminals and propose control actions to achieve the sequence;

(c) Recommend solutions for conflict situations.

The computer will be programmed to accomplish flow control on a center and later nationwide basis. Here the computer will be employed to assess future traffic flow and density by sectors and terminals and to provide the capability necessary to enable more precise flow control planning.

INDEX OF NAMES

Heavy type indicates the authorship of a chapter, no references to the author in the text of his chapter are indexed.

Adler, R. 172
Alexander, F. 12, 19
Allen, J. L. 414
Ashworth, P. V. 523
Atlas, D. 112, 116, 117, 139
Aulock, W. H. 415
Austin, P. M. 112, 128, 139

Bartnoff, S. 139
Bathker, D. 393
Bauer, P. A. 171
Baumgartner, Walter S. 433, 465
Bay, Z. 34
Bayes, 54
Beck, G. E. 233
Bell, C. 19
Benjamin, A. 172
Berger, F. B. 226, 228, 231, 234
Bigelow, Gus F. **281,** 305
Blackman, R. R. 34
Blass, J. 414
Blommendaal, R. 394
Borchardt, H. 138
Bosman, D. 279
Botwin, L. 171
Boyenval, E. H. **35,** 45
Brandon, P. S. 35, 43, **485,** 514
Bradsell, P. **141**
Brennan, L. E. 12, 19
Brown, A. 583
Brown, R. K. 265
Butler, J. 414
Bussgang, J. J. 14, 19

Carpenter, R. L. 34
Carpentier, Michel H. **1,** 6, 174
Cashen, E. R. 483
Cassegrain, W. 376
Clegg, J. E. 226, 234
Cutler, B. 483
Cutler, C. C. 394
Cutrona, L. J. 91, 95

Dauksher, W. J. 172
Davies, D. E. N. **417**
Dennis, A. S. 139
De Witt, J. H. Jr. 34
Diab, Moses A. **397,** 417
Dishington, R. H. 12, 19
Djinis, W. 305
Domville, A. R. **97,** 109, 279
Dukes, E. F. 45

Easterling, Mahlon **307,** 326

Fitch, J. L. 172
Fleischer, A. 112, 120, 122, 139
Fogy, W. **111,** 139
Foldes, P. 377, 394
Fortini, M. M. 172
Fouché, Y. 45
Fowle, E. N. 76, 95
Fried, W. R. 218, 234, 265

Gabor, D. 76, 95
Galloway, W. C. 172
Glegg, K. C. M. 265
Goldstein, R. M. **21,** 34
Golomb, S. W. 34, 326
Gray, Trevor **217,** 231, 233, 234
Gundlagh, F. W. 525

Hair, H. A. **473,** 483
Hall, G. O. 95
Hannan, P. W. 376, 394
Harris, J. 19
Hersch, W. 414
Highcock, J. **185,** 201
Hilst, G. R. 112, 139
Hitchfeld, W. 112, 123, 139
Holdberg, D. E. 306
Hogg, D. C. 387, 395
Hutson, A. R. 172

Jensen, A. S. 172
Jungfer, H. 525

Keall, O. E. **485,** 498
Kelleher, K. S. 415
Kelly, E. J. 95
Keonsian, E. 523
Kerr, D. E. 139
Killick, E. A. **417**
Komolos, S. G. 377, 394
Kotelnikow, V. A. 34
Ksienski, A. 414

Lees, R. J. 6, 45
Leith, E. N. 95
Leth-Espensen, L. **541**
Levey, L. 474, 483
Levy, G. S. 380, 393, 395
Lhermitte, R. M. 112, 122, 139
Loose, S. 138
Lowe, R. 414
Lund, C. O. **267,** 279

MacFarlane, A. G. J. 172
Mallett, J. D. 12, 19
Maggio, Thomas **397,** 414, 417
Malling, L. R. 172
Marcum, J. I. 14, 18, 19
Marshall, J. S. 139
Martin, F. W. 112, 139
Mathison, R. P. 326
Matt, Sol **203,** 214, 215
Maximon, L. C. 19
May, J. R., Jr. **172,** 514
McCrea, F. 393
McKee, D. A. 172
Miedema, H. 514
Miller, V. L. 305
Milnarich, P., Jr. 306
Mintzer, A. I. 172
Moran, M. J. 231, 234
Morgan, M. 279
Mortley, W. S. **485,** 505, 514
Mould, A. J. 265
Muhleman, D. O. 34
Murphy, J. 6

Nesbeda, P. 19
Neyman-Pearson, 54
Nicholls, N. S. **467**
Nixon, D. 393

Oakley, B. W. **565,** 583
Orford, R. J. **9,** 19
Ogg, F., Jr. 414
Oulsnam, G. 166

Paul, R. H. 306
Perkins, G. O. 305
Petrie, R. 393
Pertman, S. E. 162, 172
Pettengill, G. H. 34
Phillips, C. S. E. **253**
Pilié, R. J. 139
Potter, P. D. **373,** 377, 393, 394, 395
Potter, R. S. 395
Pound, R. V. 144, 172
Prager, R. H. 95
Prior, J. R. 265
Probert, Jones, Jr. 139

Reboul, L. **47**
Rechtin, E. 6, 371, 393, 583
Reinhardt, M. 138
Rey, T. J. 172
Ridenour, L. N. 171, 184
Rihaczek, A. W. **75,** 95
Rogers, R. R. 112, 139
Rondinelli, L. A. 414
Ruina, J. P. 19
Rumsey, H. C. 34
Rusch, W. V. T. 377, 394

Salomon, J. 201
Safran, H. 19
Saunders, W. K. 265
Schlachta, K. V. **525**
Schuster, Danver 380, 393, 394, 395
Schwartz, Mischa 19
Scolnik, M. I. 139
Sharp, E. 415
Sheehan, J. A. 95
Shelton, J. P., Jr. 415
Sherwin, C. W. 172
Shoemaker, L. E. **585**
Shnitkin, J. 415
Siebert, W. M. 76, 95
Siegert, A. J. F. 112, 139
Silber, L. M. 474, 483
Silver, S. 394, 395
Simpson, M. 414
Skolnik, M. I. 19, 265
Slobodin, L. 514
Smart, A. 265
Smith, N. J. 138, 214, 215
Solomon, K. 172
Splechtna, Dr. 138
Springer, H. **525**
Stephenson, J. G. 172
Stevens, R. 34, 326, 393
Stewart, J. L. 95
Stodola, E. K. 34
Storz, W. **525**
Stone, M. L. 139
Swerling, Peter 14, 15, 19

Thomson, J. H. 34
Thorne, T. G. 226, 228, 234
Thrane, J. 465
Titsworth, R. C. 326
Tukey, J. W. 34
Tuller, W. G. 172

Van Den Broek, G. **173**
Van, Voorhis 172
Vleck, J. H. 395
Victor, W. K. 34, 326, 393
Viterbi, A. 326

VIVIAN, W. E. 95
VOGEL, M. 123
VOSS, R. A. 306

WALLACE, P. R. 112, 139
WEISS, H. G. **327**
WEISS, J. E. 306, 414
WESLEY, A. S. 109
WESTERFIELD, E. C. 76, 95
WHITE, J. E. 306
WHITE, W. O. 172
WILKERSON, J. R. 172
WILMER, R. W. 6, 215
WIRTH, W. D. **525**
WOOD, J. 265
WOODWARD, J. E. **235**
WOODWARD, P. M. 53, 61, 65, 76, 77, 95, 174

YOUNG, A. P. **515,** 523

ZAFFARANO, F. P. 172

INDEX OF PLACES AND INSTITUTIONS

Academy of Technical Sciences Copenhagen 267
Admiralty Surface Weapon Establishment 417
Advanced Research Projects Agency 405
Aerospace Corporation 75
Airborne Instruments Laboratory 235

Bell Laboratories 35
Bendix Corporation 405
Birmingham University 417
Britain 203

Canberra, Australia 433
C.F.T.H. Bagneax (Seine) France 1
C.N.E.T. 47
Copenhagen 267
Crimean Station 22
Cryenco Inc. 368

Decca Navigator, Company Limited 233
Decca Radar, Ltd. 141, 151, 166, 217, 233
Deep Space Instrumentation Facility 373, 386
Directions des Recherches et Moyens d'Essais 183
DVL Institute of Atmospheric Physics 138

El Adem 105
Energy Systems Inc. 360
England 203
Europe 19
Exmoor 104

Federal Aviation Agency, Washington D.C. 585
Ferranti Ltd., Edinburgh 265
Flugwissenschaftliche Forshungsanstalt, München 138
France 183

General Electric Co. 97, 203, 473
Goldstone, Calif. 311, 433, 443

Hahn-Meitner-Institut (Berlin) 525
Hartlepool 199
Heinrich-Hertz Institut (Berlin) 525
Hilger and Watts Ltd., London 514
Hungary 21

Institut fur Flugfunk und Mikrowellen 111
Institut fur Funk und Mathematik, Germany 525

Jet Propulsion Laboratory, Pasadena 21, 22, 33, 34, 307, 308, 313, 319, 324, 371, 373, 376, 377, 386, 389, 392, 393, 433, 437, 443, 463, 464, 465, 583
Jodrell Bank 22
Johannesburg, South Africa 433

Keuffel & Esser of Hoboken, New Jersey 336

Laboratoire Central de Télécommunications 173, 183, 184
Lamesa, California 339
Libyan Desert 107
Lincoln Laboratory 327, 339, 340, 341, 346, 408, 473

Madrid, Spain 433
Malvern 578, 579
Malvern Hills 201
Marcel Dassault 182
Marconi Company 35, 37, 42, 485, 514, 515
Massachusetts Institute of Technology 22, 339, 371 (see Lincoln Lab.), 473
Maxson Electronics 410
Middlesbrough 199
Millstone Hill Station, Westford 327
Microwave Laboratory, Copenhagen 267
Mount Tiefort 384, 385

National Aeronautics & Space Administration 373, 433
National Aviation Facilities Experimental Center at Atlantic City, New Jersey 248, 594
New England 329
North American Aviation 336, 340

Pacific 204
Pasadena, California 434

RADC 410
Radiation Research Laboratory, Stanford 360
Rand Corporation 19
Rantec Inc. 366
Raufoss, Ammunisjonsfabrikker, Norway 267, 278

Raytheon Company 366
Rugby 37
Rome Air Development Center, U.S.A. 397, 405
Royal Radar Establishment 35, 37, 185, 201, 253, 265, 467, 565

SHAPE, Technical Centre, The Hague, Holland 541, 578
Southampton University 171

Terma, Electronic Industry 267
T.R.W. Space Technology Laboratories 9

United Kingdom 85
United States 19, 35, 183, 417
U.S. Signal Corps. 21

Varian Associates 363

West Midlands 200
White Sands Missile Range (WSMR) 281, 305
Wolverhampton 105
Worcester 104
Woomera 433

INDEX OF SUBJECTS

Adaptive filter 52, 54, 175
Aerial, conical beam 228
Aerial, drift stabilised 227
Aerial, fixed azimuth 227
Aerial, horn 188
Aerial, squint 197
Aerial, stripline construction 232
Airborne early warning system 203
Angels 111
ARINC/540 229
Area moving target indication 165
Avenger, TBM 203

Balanced phase detector 146
Barium titanate transducer 144

Cathode ray tube characteristics 194
Center-metroplex 586, 590, 594
Ceramic transducers 144
CHIRP 3, 35, 45, 82, 86
Clear air turbulence 111
Cloud 111
Clutter lock 155, 156
Clutter rejection 569, 577
Combat information centre 204
Combiner tape 555
Correlation detection 51, 52
Correlation detection = nt to adaptive filter 52

Delay line 181
Detection probability 11
Digital storage 165
Digitalisation 525
Dispersive networks 485
Display storage tube 165
Doppler navigation 217
Doppler navigation calibration 232
Double cancellation 148
Ducting 207

Earth 10, 25, 433

Ground refectivity 97, 107, 214

Hawkeye, E-2A 204
Homodyne receiver 226
Hopping strobe 520

Inner flight number 532
i.f. cancellation 155, 156, 160

Jupiter 21, 22

Lambert's law 99

Marcov algorithm 532
Mars 21, 22
Mercury 21, 22
Microcircuits 207
Microwave transmitting valve 467
Modulators, network type 469
Moon 5, 6, 21

Nagakami distribution 114, 120
Networks, delay response 503
Networks, dispersive 485
Neyman-Pearson criterion 54
Non-cooperative target 9

Optimum waveform 83

Phase shifter, latching 474
Photographic recorder 186
Piezo-electric grating 506
Plan position indicator 516
p.r.f., staggered 149, 166
Pulse compression, see CHIRP 35, 75, 82, 161, 575
Pulse Doppler 3, 162, 209

Quartz delay line 142, 143

Radar, APS-96 204
Radar clutter, see pulse Doppler and radar moving target indication 209
Radar, early airborne 203
Radar, FM 82, 152, 225
Radar, IFF 206
Radar, ground mapping 75, 91, 196
Radar K_a-band 111
Radar, L-band 141, 144, 158
Radar, low flying aircraft 173, 183
Radar, moving target 141, 153, 157, 158, 159, 160, 161, 165, 166, 173, 209
Radar, resolution 194
Radar, S-band 97, 141, 144, 158
Radar, sideways looking 185
Radar, synthetic aperture 160
Radar, X-band 111, 144, 189
Radechon 165
Radome 229
Rayleigh surface waves 510
Reflecting grating 507
Rice distribution 114

Satellite, 1957 Beta 18, 19

Schwartz inequality 67, 78
Sea clutter 210, 214
Sea state 214, 232
Sief dune 107
Signal clutter, variation with altitude 214
Signal generation by frequency multiplication 231
Snow, reflectivity 105
Stalo 144, 158
Stripline 232
Swerling, Case I 9, 12, 13, 17, 19
Swerling, Case III 17, 19
Synthetic PPI 516

Target analysis by ear 180
Terabil 107
Terminal areas 585
Thyratron 147, 468
Time jitter 147, 158
Track, apparent 546
real 546
Tracker ball 518
Tracking in the turn 574
Transistors 145
Travelling wave tube 471

Ultrasonic delay line 142, 143
Ultrasonic dispersive grating 506

Valve, microwave transmitting 467
Valve, travelling wave tube 471
Venus 21, 22, 24, 25, 26, 32, 33
Velocity shaping 148
Video integration 190

Water reflectivity 101